

ABSTRACT

Normativity and Mathematics: A Wittgensteinian Approach to the Study of Number

J. Robert Loftis

I argue for the Wittgensteinian thesis that mathematical statements are expressions of norms, rather than descriptions of the world. An expression of a norm is a statement like a promise or a New Year's resolution, which says that someone is committed or entitled to a certain line of action. An expression of a norm is *not* a mere description of a regularity of human behavior, nor is it merely a descriptive statement which happens to entail a norm. The view can be thought of as a sort of logicism for the logical expressivist—a person who believes that the purpose of logical language (sentential connectives, quantifiers, etc.) is to make explicit commitments and entitlements that are implicit in ordinary practice. The thesis that mathematical statements are expression of norms is a kind of logicism, not because it says that mathematics can be reduced to logic, but because it says that mathematical statements play the same role as logical statements.

I contrast my position with two sets of views, an empiricist view, which says that mathematical knowledge is acquired and justified through experience, and a cluster of nativist and apriorist views, which say that mathematical knowledge is either hardwired into the human brain, or justified a priori, or both. To develop the empiricist view, I look at the work of Kitcher and Mill, arguing that although their ideas can withstand the criticisms brought against empiricism by Frege and others, they cannot reply to a version of the critique brought by Wittgenstein in the *Remarks on the Foundations of Mathematics*. To develop the nativist and apriorist views, I look at the work of contemporary developmental psychologists, like Gelman and Gallistel and Karen Wynn,

as well as the work of philosophers who advocate the existence of a mathematical intuition, such as Kant, Husserl, and Parsons. After clarifying the definitions of “innate” and “a priori,” I argue that the mechanisms proposed by the nativists cannot bring knowledge, and the existence of the mechanisms proposed by the apriorists is not supported by the arguments they give.

NORTHWESTERN UNIVERSITY

**Normativity and Mathematics:
A Wittgensteinian Approach to
the Study of Number**

A DISSERTATION

**SUBMITTED TO THE GRADUATE SCHOOL
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS**

for the degree

DOCTOR OF PHILOSOPHY

Field of Philosophy

By

J. Robert Loftis

EVANSTON, ILLINOIS

December 1999

© Copyright by J. Robert Loftis 1999
All Rights Reserved

ABSTRACT

Normativity and Mathematics: A Wittgensteinian Approach to the Study of Number

J. Robert Loftis

I argue for the Wittgensteinian thesis that mathematical statements are expressions of norms, rather than descriptions of the world. An expression of a norm is a statement like a promise or a New Year's resolution, which says that someone is committed or entitled to a certain line of action. A expression of a norm is *not* a mere description of a regularity of human behavior, nor is it merely a descriptive statement which happens to entail a norms. The view can be thought of as a sort of logicism for the logical expressivist—a person who believes that the purpose of logical language (sentential connectives, quantifiers, etc.) is to make explicit commitments and entitlements that are implicit in ordinary practice. The thesis that mathematical statements are expression of norms is a kind of logicism, not because it says that mathematics can be reduced to logic, but because it says that mathematical statements play the same role as logical statements.

I contrast my position with two sets of views, an empiricist view, which says that mathematical knowledge is acquired and justified through experience, and a cluster of nativist and apriorist views, which say that mathematical knowledge is either hardwired into the human brain, or justified a priori, or both. To develop the empiricist view, I look at the work of Kitcher and Mill, arguing that although their ideas can withstand the

criticisms brought against empiricism by Frege and others, they cannot reply to a version of the critique brought by Wittgenstein in the *Remarks on the Foundations of Mathematics*. To develop the nativist and apriorist views, I look at the work of contemporary developmental psychologists, like Gelman and Gallistel and Karen Wynn, as well as the work of philosophers who advocate the existence of a mathematical intuition, such as Kant, Husserl, and Parsons. After clarifying the definitions of “innate” and “a priori,” I argue that the mechanisms proposed by the nativists cannot bring knowledge, and the existence of the mechanisms proposed by the apriorists is not supported by the arguments they give.

Acknowledgements

Molly Hinshaw, who normally gets paid good money to edit academic manuscripts, agreed to look this one over for free. For this and many other generousities, I thank her. I would also like to thank my committee, Arthur Fine, Meredith Williams, and Michael Williams, for their helpful comments and for working with me. I owe a debt of gratitude to the support staff at Northwestern, especially Donna Choccol, for making the years spent in the philosophy department there much more bearable. Finally, I thank my parents for their love and support.

Bibliographic Conventions

Most works are referred to by the author's name, date of publication, and page number. Reprinted works will have two publication dates; the page number is taken from the later edition. Multivolume works include the volume number as a Roman numeral and the page number as an Arabic numeral.

Most locations in Wittgenstein's works are identified by the section number. Citations of middle period works, which often feature long sections, include the number of the paragraph within the section enclosed in square brackets after the section number. When a page number is used in a citation of Wittgenstein's work, I have explicitly included the abbreviation "p." or "pp."

Wittgenstein's works are referred to using the common abbreviations for their titles:

<i>AWL</i>	<i>Wittgenstein's Lectures, Cambridge 1932–1935</i> (Wittgenstein 1979c)
<i>BB</i>	<i>The Blue and Brown Books</i> (Wittgenstein 1969b)
<i>CV</i>	<i>Culture and Value</i> (Wittgenstein 1980)
<i>NB</i>	<i>Notebooks, 1914–1916</i> (Wittgenstein 1961)
<i>OC</i>	<i>On Certainty</i> (Wittgenstein 1969a)
<i>PI</i>	<i>Philosophical Investigations</i> (Wittgenstein 1967a)
<i>PG</i>	<i>Philosophical Grammar</i> (Wittgenstein 1978)
<i>PR</i>	<i>Philosophical Remarks</i> (Wittgenstein 1975)
<i>RFM</i>	<i>Remarks on the Foundations of Mathematics</i> (Wittgenstein 1979a)
<i>WWK</i>	<i>Wittgenstein and the Vienna Circle</i> (Wittgenstein 1979b)
<i>Z</i>	<i>Zettel</i> (Wittgenstein 1967b)

Classical and very famous works are also referred to by their abbreviations

<i>TT</i>	Boethius, <i>The Theological Tractates</i> (c.510/1936)
<i>CPR</i>	Immanuel Kant, <i>Kritik der Reinen Vernunft</i> (1781 and 1787/1969)
<i>Ethics</i>	Baruch Spinoza, <i>Ethica Ordine Geometrico Demonstrata</i> (1677/1992)

Table of Contents

Abstract.....	iii
Acknowledgements.....	v
Bibliographic Conventions.....	vi
Introduction.....	1
Chapter 1: Empiricism.....	8
<i>Introduction</i>	8
<i>Psychologism</i>	10
<i>A Priori versus Empirical</i>	28
<i>Analytic versus Synthetic</i>	31
<i>How Statements of Basic of Basic Arithmetic and Set Theory Are Known</i>	45
<i>Ontology</i>	53
<i>History</i>	58
Chapter 2: Critique of Empiricism.....	65
<i>Introduction</i>	65
<i>The Middle Period Argument</i>	71
<i>The Late Period Argument</i>	82
<i>The Historical Objection</i>	112
<i>The Application Problem</i>	119
<i>Conclusion</i>	127
Chapter 3: Nativism and Apriorism.....	129
<i>Introduction</i>	129
<i>Two Unworkable Definitions of ‘Innate’</i>	134
<i>The Representationalist Nativist Position</i>	144
<i>The Intuitionist Apriorist Position</i>	174
<i>The Combined Nativist-Apriorist Position</i>	197
Chapter 4: Critique of Nativism and Apriorism.....	221
<i>Critique of Representationalist Nativism</i>	221
<i>Critique of Intuitionist Apriorism</i>	234
<i>Critique of Dispositionalist Nativism and Negative Apriorism</i>	243
Chapter 5: Numbers as Normative Facts.....	258
<i>Introduction</i>	258
<i>Are Norms Inherently Social?</i>	260
<i>The Objectivity of the Expression of Norms</i>	285
<i>The Universality of Mathematical Norms</i>	299
Reference List.....	316

Introduction

In what follows I will argue for the Wittgensteinian thesis that mathematical statements are expressions of norms rather than descriptions of the world. An expression of a norm is a claim that someone is committed or entitled to a certain line of action. Promises and New Year's resolutions, statements where the speaker commits herself to an action or behavior, are one important class of expressions of norms. Another is statements like "You ought to be quiet during the movie," statements that acknowledge commitments or entitlements that the speaker already believes to be in force. An expression of a norm is *not* a mere description of a regularity in human behavior, such as "India is governed by a complex caste system." An expression of a norm can name a rule that is followed nowhere in the world. Expressions of norms must also be distinguished from statements that merely entail norms. "It is raining" may imply a certain course of action involving umbrellas, but it is still clearly a descriptive statement. Finally, there is a certain minimal sense in which every assertion expresses a norm. If I say "grass is green" I am committing myself to the belief that grass is green and saying that others ought to be committed to that belief as well. The sorts of statements I am referring to as expressions of norms must express richer norms than that. "Grass ought to be green, even though my lawn remains stubbornly brown" is an expression of a norm in the sense I will be using here.

My thesis can be thought of as a sort of logicism for the logical expressivist. Logical expressivism, as outlined by Robert Brandom and others, is the belief that

the purpose of logical language (sentential connectives, quantifiers, etc.) is to make explicit commitments and entitlements that are implicit in ordinary practices. The conditional, for instance, allows us to make explicit the fact that commitment to one proposition is also commitment to another. The conditional thus allows us to express normative judgements that otherwise would have been only implicit in our practices of committing ourselves. The thesis that mathematical statements are expressions of norms is a kind of logicism, not because it says that mathematics can be reduced to logic, but because it says that mathematical statements play the same role as logical statements.

More broadly, I am arguing against the occasionally heard metaphor that mathematics is the “language of nature.” Mathematics is not the language of nature; it is the language of humans. It is about humans and the way they think they ought to relate to each other and the world. Actually, as I see it, nature does not have a language at all. Language is ineliminatively normative, and norms are the product of human interaction. However, most language can be *about* nature, and even this much is not true of mathematics.

My argument for believing that mathematical statements are normative is drawn from Ludwig Wittgenstein, principally from Book I of the *Remarks on the Foundations of Mathematics*. The remarks collected there give a series of hypothetical situations that highlight the contrast between the use of mathematical statements and the use of descriptive statements, while bringing out similarities between mathematical statements and more straightforwardly normative statements. As I reconstruct it in Chapter 2, the argument we are presented with is a form of inference to the best explanation: the best way to explain the linguistic habits and instincts brought out by the examples is to say

that mathematical statements are not descriptions but expressions of norms. This argument is strengthened in Chapter 4, when competing explanations taken from empirical psychology of our use of mathematical statements are shown to be misdirected or incoherent.

The project is organized dialectically. The two halves of my dialectic are the empiricist view of mathematics, which claims that mathematical knowledge is learned from experience and justified by experience, and a cluster of nativist and apriorist views, which claim that mathematical knowledge is either hardwired into the human brain or justified a priori or both. I argue that both wings of this dialectic are mistaken on Wittgensteinian grounds—they fail to account for the normative role mathematical statements play in both ordinary and scientific language. The view that mathematical statements are expressions of norms then appears as the synthesis of the dialectic—it is what we are left with after the nature of the original debate is cleared up. Because of his metaphilosophical views, Wittgenstein was loath to present anything one might call positive theorizing. As a result, the synthetic portion of my dialectic will go beyond Wittgenstein's own ideas. My chief goal in this section will be to show that mathematics retains all the objectivity one expects of it, even once it is seen as a normative discipline.

The first chapter draws on the work of John Stuart Mill and Philip Kitcher to develop as strong a case for mathematical empiricism as possible. I argue that the empiricist position can withstand the objections brought against it by Gottlob Frege and his successors. I also discuss the axiomatic system Kitcher developed for empiricist arithmetic and attempt to flesh out some of the claims Kitcher made about the historical

genesis of mathematics. The second chapter presents the Wittgensteinian criticism of empiricism, including inference to the best explanation argument drawn out of Book I of the *Remarks on the Foundations of Mathematics*. This argument has often been discounted by commentators who conflate it with earlier, inferior arguments from the *Philosophical Remarks* and the *Philosophical Grammar*. Distinguishing these arguments is an important part of the task of Chapter 2.

Chapter 3 attempts to give the strongest possible arguments for the nativist and apriorist positions, just as Chapter 1 attempted to give the strongest possible argument for the empiricist position. Most of the nativist arguments will draw on contemporary cognitive and developmental psychology, while the apriorist arguments will come from those writing in the tradition of Immanuel Kant, Edmund Husserl, and the positivists. Here a major task will simply be clarifying the definitions of “a priori” and “innate.” I will argue that a couple of the definitions used in the psychological literature are simply unworkable. I wind up recognizing four workable definitions: one strong and one weak definition each for the terms “innate” and “a priori.” As it turns out, the two weak definitions share a modal structure—they are both about what would happen in the normal development of a person with normal cognitive abilities—and are therefore co-extensive. This leaves three points of view to defend: the strong nativist, the strong apriorist, and the weak combined position. As a part of this defense, I survey the empirical results from cognitive developmental psychology and neuropsychology, including studies of the arithmetic abilities of animals, infants, children, and normal and brain-damaged adults, as well as research using EEG, CT, and MRI technology.

Chapter 4 turns around and presents a Wittgensteinian critique of these ideas. In my treatment of nativism, I make no effort to challenge the empirical data. Nor do I challenge the psychological models of this data. Instead, I attack the philosophical interpretation given to the data. The strong nativist position asserts the existence of certain innate representations, and the psychological models of the empirical data are taken to show the existence of just this kind of innate representation. I argue that even if these models were to be true, the structures identified could not bring knowledge, because they can never be wrong. They are always representing some mathematical structure correctly. The strong apriorist position, on the other hand, asserts the existence of a priori intuitions. Against this view, I claim that all of the arguments that are meant to show the existence of such an intuition actually promote a Wittgensteinian idea: that there is a way of acting that is not an interpretation, where one acts without justification, but not without right. The final position I have to confront is the weak combined nativist/apriorist view. Here I change tactics, and simply accept the conclusions of the nativist/apriorist. The definitions of “innate” and “a priori” are now so weak that there is no contradiction in saying both that mathematics is innate and a priori, and that mathematical statements are expressions of norms.

In the final chapter I turn to the task of presenting a positive account of mathematics. My focus will be on the objectivity of mathematics. Expressions of norms have the same kind of objectivity that descriptive statements have. Like descriptive statements, normative statements can be challenged and defended. There is also a sense in which a universally held normative statement is wrong. Finally, normative statements are capable

of de re ascription. They are thus about objects that exist apart from language. The first task in arguing for these claim will be to clarify the social nature of normativity. Many people have argued that Wittgenstein's discussion of rule following implies that one can only obey rules if one is a part of a community of rule followers. Others argue just as strenuously that Wittgenstein believed no such thing. In the first part of Chapter 5, I will argue that both sides are mistaken. Wittgenstein's thoughts on rules are not sufficient to either mandate a social conception of norms or rule it out. Our understanding of the sociality of norms must come from another source. I will argue that norms are weakly social. Norms must be seen against the background of a practice. This practice can in principle be individual; however, as a matter of fact, all of the practices that sit behind real norms are social. Once the social nature of norms is understood, it will be possible to see their objectivity. Drawing Brandom's writings, I will argue that the objectivity of our language use comes from our practice of deontic scorekeeping. This practice applies both to descriptive statements and to expressions of norms. Once the objectivity of norms in general has been established, I will take some time to discuss the objectivity of mathematical norms in particular. I claim that any language that features singular terms must also allow those singular terms to be counted. Thus even if the language does not contain mathematical elements, it must be compatible with the introduction of such elements. This argument will be an extension of Brandom's argument that any language that is compatible with the introduction of basic sentential connectives must also be compatible with the introduction of singular terms. By arguing this I hope to assuage

fears that Wittgenstein's approach to mathematical statements implies that they are somehow culturally relative.

Empiricism

Introduction

The motive for empiricism in the philosophy of mathematics is the conviction that the true nature of mathematical knowledge lies in the simple activities children perform when they learn arithmetic—collecting marbles, moving pieces on a game board, etc. These activities might be significant for any number of reasons. They might have an epistemological function, serving to legitimate belief in mathematical statements; they might have a causal function, serving as the learning process that brings about mathematical knowledge; or their role might be ontological and semantic, serving as the objects referred to by mathematical statements. When John Stuart Mill introduced the modern empiricist account of mathematics, he endorsed empiricism along all of these dimensions.¹ Mill's arguments were not always very strong, and his exposition was marred by technical mistakes. These problems drew criticism from many, as well as outright ridicule from Gottlob Frege, and as a result mathematical empiricism was not taken seriously in philosophy for the bulk of the twentieth century.

Nevertheless, recent trends in philosophy have led some to reconsider empiricist ideas about mathematics. W. V. O. Quine's argument in "Two Dogmas of Empiricism" (1951/1961) that all knowledge is subject to revision in light of experience

¹ See Mill (1843) esp. Book 2 chapters 5 and 6 and Book 3 chapter 24.

has led to a retreat from apriorist accounts of any form of knowledge. Within the philosophy of mathematics, many who had in the past endorsed Platonist theories have embraced empiricist ideas in response to Paul Benacerraf's paper "Mathematical Truth" (1973). In that paper, Benacerraf posed a question for Platonists: If mathematical objects are ideal beings outside of space and time, how can humans, whose knowledge comes through spatial and temporal channels, ever have knowledge of them? (The original challenge was phrased in terms of a causal theory of knowledge; however, it has since been refined to be amenable to a variety of epistemologies.²) The response has largely been to redescribe mathematical objects so that we can have access to them through traditional sensory channels. (See Penelope Maddy 1991 for descriptions of this trend.) Maddy herself (1990) has proposed a theory in which some mathematical objects are directly perceived. Others, such as Resnick (1997) and Shapiro (1983), have proposed theories in which the justification for mathematical beliefs is bound closely to ordinary perception, in part in response to Benacerraf's challenge.

The contemporary thinker who has done the most to revive empiricist ideas in the philosophy of mathematics is Philip Kitcher. Kitcher's book *The Nature of Mathematical Knowledge* is the first rigorous exposition of the ideas Mill was reaching for. In this chapter, I would like to argue that Kitcher's account can withstand objections typically

² See for example Field (1989, 25–30) and Maddy (1990, ch. 2 §1). Steiner (1975, ch. 4), on the other hand, argues that Benacerraf-like challenges depend on an unacceptable form of the causal theory of knowledge.

brought against empiricism by Frege and the anti-empiricists of the twentieth century. Let's start by looking at the epistemological background that Kitcher sets up for his theory.

Psychologism

When Mill gave his account of the empirical origin of mathematics, he took it for granted that he could simultaneously describe the causal origin and epistemic justification of mathematical belief. Kitcher, too, wishes to unite causality and epistemology; however, he gives a much more explicit account of the bond between these two dimensions of empiricism. Kitcher describes his work as a piece of *psychologistic* epistemology, which he defines as a theory that looks at the means by which one actually acquires a belief in order to determine if that belief counts as knowledge (Kitcher 1984, 13 ff.). An *apsychologistic* epistemology, on the other hand, looks only to the relation of a belief to other beliefs to determine whether that belief counts as knowledge, ignoring the way it was acquired. Now this distinction between psychologistic and apsychologistic epistemology can be construed in two ways. On the one hand, psychologism could be seen as a theoretic claim about the proper analysis of knowledge, namely, that it must refer in some way to the cause of the belief that is supposed to count as knowledge. Apsychologistic epistemology would then be the claim that the proper analysis of knowledge only refers to the beliefs an individual holds and the relationships between them. On the other hand, psychologism may be seen as a metatheoretic claim about the

best way to develop an analysis of knowledge, namely, that it involves an empirical investigation into the way people actually come to believe different things. The metatheoretic claim does not imply the theoretic claim, nor does the theoretic claim imply the metatheoretic claim. We may, at the end of our empirical investigation, decide that the conditions under which beliefs are formed are so heterogeneous that they cannot serve as part of our analysis of knowledge. On the other hand, we might endorse the theoretic claim without actually having engaged in an empirical investigation of the origin of belief. We could include in our analysis of knowledge a reference to the causes of belief in general without having to do any empirical work as to what kinds of causes there are or how they interact with knowledge. This is in fact what Kitcher does, so in what follows I will treat psychologism solely as a thesis about what belongs in our analysis of knowledge.

Kitcher's most thorough treatment of psychologistic epistemology is his 1992 essay "The Naturalists Return." There he states in general form the argument that he used to defend psychologism in *The Nature of Mathematical Knowledge* (1984, 15 ff.), an argument that Kitcher also sees in Harman (1973, ch. 2), Goldman (1979), and Kornblith (1980). In its baldest form, this line of reasoning asks us to suppose that a person knows P, knows that $P \rightarrow Q$, and moreover believes that Q. She, however, does not believe that Q because she has put together the facts that P and $P \rightarrow Q$. Rather, she has a hunch, or is taking the word of an unreliable authority. (In *The Nature of Mathematical Knowledge* Kitcher suggests a mathematical belief accepted because of a dream, trance, or "fit of

Pythagorean ecstasy.”) If we only take features of the propositions believed into account in our epistemology, we will be forced to say that such a person knows Q. The statement Q stands in the right relation to her other beliefs. Her ignorance of this is an accidental feature about her, which should not affect the justification of the statement Q. But the person in our example clearly does not know Q. Therefore, an analysis of knowledge that does not refer to the circumstances under which belief is formed does not capture everything there is to knowledge.

A couple points are worth noting about this argument. First, although it has its roots in Edmund Gettier’s famous argument in “Is Knowledge Justified True Belief?” (1963/1986), it is not quite the same. Gettier’s examples all involved people who have a true belief that they have arrived at by a means that is usually reliable, but in this case has failed them. For instance, suppose Bob believes that a colleague owns a Ford because he has seen her driving a Ford. This kind of inference is usually reliable, since people generally drive their own cars. In this case, however, the colleague was driving a friend’s car. Nevertheless she does happen to own a Ford, so Bob’s belief is correct. Cases like this were supposed to show that knowledge is not merely justified true belief. Bob’s belief about his colleague were justified and true, yet he did not have knowledge. Now Kitcher’s example, unlike Gettier’s, does not involve a person using a process that is generally reliable. As a result it changes the point of the argument in subtle and important ways. Kitcher’s argument is designed to show the need for considering belief-forming processes in general, but Gettier’s argument is directed at a view that is already looking at

belief-forming processes. That is why the examples used involve processes that are generally reliable which happen to have failed in this case. The justification that Gettier says is not a sufficient condition to turn true belief into knowledge could be either a psychologistic or an apsychothetic justification. Therefore the force of Gettier's arguments needs to be separated from the force of Kitcher's.

The second point is that Kitcher's argument only shows that there is a performance aspect to justification.³ Justification is a process one can engage in or fail to engage in. The argument shows that one must actually perform a justification to be a knower. It does not show, however, that this performance aspect is inseparable from the other aspects of knowledge, so that one cannot understand knowledge without understanding how justifications are performed. Therefore Kitcher has not shown that apsychothetic definitions of knowledge are incoherent, merely that they are incomplete. In fact, it may be the case that all one needs to add to an apsychothetic account of knowledge is the directive, "Now actually do the things outlined in this account." Nevertheless, it is clear from this argument that an analysis of knowledge that includes an account of belief-forming processes captures more of knowledge than one that does not, and this is sufficient to motivate Kitcher's project.

The psychologistic analysis of knowledge that Kitcher proposes runs like this:

X knows that p if and only if p and X believes that p and X's belief that p was produced by a process which is a warrant for it (1984, 17).

³ I owe this point to Meredith Williams and Michael Williams.

A process is a warrant for a belief if it produces the belief in “the right way,” that is, the way to acquire the belief so that it counts as knowledge. Obviously, for this analysis to be at all informative, just what the “right way” is must be spelled out. This is not a project Kitcher undertakes in *The Nature of Mathematical Knowledge*. In that work he only introduces the general analysis of knowledge to codify his psychologism before starting on his account of mathematical knowledge. However, in many other essays, such as “The Naturalists Return,” he indicates strong sympathy for some sort of reliabilism, the theory that the “right sort” of belief-generating process is one that can be counted on to yield truth.

By beginning his statement of mathematical empiricism by declaring himself a psychologistic epistemologist, Kitcher is compounding his sins against the dominant movements of the twentieth century. For a very long time, opposition to psychologism was so unified that the word itself had a pejorative meaning. The sort of views that the word referred to, however, was not often clearly stated. Nicola Abbagnano (1967), writing for the *Encyclopedia of Philosophy*, places the earliest use of “psychologism” in early nineteenth century Germany, where it referred to an anti-Hegelian movement led by Jakob Friedrich Fries and Friedrich Beneke that held that introspection was the only route to philosophic truth. In the mid-nineteenth century John Stuart Mill endorsed psychologism in a more general sense by classifying logic as a branch of psychology (1865/1973, 145–46). The first notable opposition to psychologism came from Gottlob Frege, who was issuing screeds against psychologism at the same time he was firing off

the denunciations that would drive empiricism out of academic philosophy of mathematics. The mid- and late-nineteenth century writers Frege was criticizing were a heterogeneous lot. Edmund Husserl's *Philosophie der Arithmetik* (1891/1970) attracted Frege's ire by claiming that number could not be defined, but instead should be associated with irreducible features of perception. Benno Erdmann's *Logik* (1892) was found guilty of equating truth with widespread acceptance. Frege's polemics were largely ignored in his lifetime, the most prominent opposition to psychologism coming from neo-Kantians of the Southwest and Marburg schools. Paul Natorp, for instance, criticized those who held, "that the true grounding of knowledge be sought in [its] relationship to the subject, in subjective 'consciousness'" (1887/1981, 248–49). Under such a view, "logic becomes unavoidably dependent on psychology, which conclusion at least the most consistent advocates of the subjective viewpoint have not shied away from" (ibid.). Natorp challenges this view on the grounds that a logic based on empirical psychology cannot have the "fundamental validity" required of it. "A science which according to its name and claim treats knowledge in general and its laws may not be dependent for its own grounding on a particular scientific knowledge" (ibid., 151). By 1896, Edmund Husserl, too, had come around to the idea that psychologism was a danger to the apodeicity of knowledge. Disavowing the approach he took in *Philosophie der Arithmetik*, he wrote the *Prolegomena zur reinen Logik*, the extensive, propaedeutic first volume of his seminal *Logische Untersuchungen* (1900–1/1970). The aim of the *Prolegomena* was solely to criticize psychologism, which he took to be the thesis that

logic was a “technology” of thought based on empirical psychology. Frege’s attack on *Philosophie der Arithmetik*, as well as the related letters which were exchanged between the two thinkers, is often taken to be the source of Husserl’s conversion. However J. N. Mohanty (1974, 1980) has argued that Husserl’s ideas developed before he encountered Frege. Frege was certainly read closely by Russell and Wittgenstein, who worked antipsychologism into the fabric of logical positivism. But for too many of the logical positivists, the word “psychologism,” like the word “metaphysics” was more a blanket epithet for anything they opposed than a coherent doctrine they were opposing. What emerges from these shifting definitions is the picture of psychologism as the view that the investigation of human knowledge, whether it be conducted under the rubric of “epistemology” or “logic,” should be an empirical science. Kitcher’s psychologism clearly belongs to this lineage—he says to study knowledge we must study the causes of belief, and the study of causes is an empirical enterprise. To see if Kitcher runs afoul of the trenchant critiques of psychologism of the past, I will examine the two thinkers who did the most to promote antipsychologism, Husserl and Frege, beginning with Frege.

Frege’s shotgun attack on psychologism often threw shrapnel at mathematical empiricism as well, but in this presentation I will try to keep the two issues separate. Michael Resnick (1980, ch. 1) identifies four trends in epistemology that seem to be comprised by Frege’s image of the psychologist: the identification of meanings or concepts with mental entities, the preference for descriptions of the genesis of mathematics over the sort of reductive definitions he offered in his own work, the

reduction of truth to acceptance, and the identification of normative laws of thought with general regularities in human thinking.⁴ The identification of meanings of concepts with mental entities is probably the central feature of the psychologism that Frege opposed. Frege's central complaint against this idea is that it renders communication impossible. If one person's concept of the number two is an idea in their head, then everyone would have their own concept of the number two "and then we should have many millions of twos on our hands" (1884/1968, §27). Moreover, if there are different twos, they might have different properties: "as new generations of children grew up, new generations of twos would continually be being born, and in the course of millennia these might evolve, for all we could tell, to such a pitch that two of them would make five." (ibid.).

The second strongest trend in the epistemologies Frege labeled psychologist is a preference for developmental descriptions over reductive definitions. Frege's epistemology of mathematics consisted entirely of defining number in terms of simpler concepts and building an axiomatic system that would yield all mathematical truths. Whether these simpler concepts actually played a role in the development of mathematical knowledge was irrelevant to Frege. The most notable example of a genetic account that Frege felt misguided was Husserl's *Philosophie der Arithmetik*. In that work, Husserl argues that number cannot be defined, specifically criticizing Frege's attempts to define number. Since number cannot be defined, Husserl asserts, we must look for its origins in experience. Husserl finds this origin in the process of abstraction: number is a

⁴ Notturmo (1985) comes up with a similar analysis of Frege's attack on psychologism.

simple concept drawn from the world by focusing on some aspects of experience and ignoring others. In his review of *Philosophie der Arithmetik* Frege seems to doubt that the process of abstraction can yield anything, sarcastically remarking, “Inattention is an exceedingly effective logical power; whence presumably, the absentmindedness of scholars.” (1894/1977, 9). More importantly, he feels that reliance on abstraction leads back to the idea that meanings are mental entities. An idea of number, once abstracted from experience, is a purely subjective entity. Yet if we are to communicate, it must be objective. Thus Frege claims that Husserl blurs the boundary between the subjective and the objective (ibid.). Now, one can easily imagine that not all developmental accounts of arithmetic would lead to the equation of meanings with mental entities in quite this way. However, Frege also has a general objection to the use of genetic accounts: they lead to relativism. By using features of the origin of a mathematical concept in accounting for our knowledge of that concept, one conflates the context of justification with the context of discovery. The message of a genetic account seems to be that the process of discovering an idea also justifies the idea. If one assumes that a process of justification must always lead to truth, one is left saying that things become true when people believe them. Something like this argument seems to be at work in Frege’s attack on Stricker (1883). Stricker presented an account of mathematics based around the origin of the number concept in muscular sensation. By the end of Frege’s attack (1884/1968, v–vi) on this idea, he is boldly declaiming, “Never let us take a description of the origin of an idea for a definition, or an account of the mental and physical conditions on which we become

conscious of a proposition for a proof of it.... Otherwise in proving Pythagoras' theorem we should be reduced to allowing for the phosphorous content of the human brain; and astronomers would hesitate to draw any conclusions about the distant past, for fear of being charged with anachronism" (ibid., vi).

In the argument above, the equation of truth and acceptance is presented as the disastrous consequence of the use of genetic accounts in epistemology. Frege also felt that it was a thesis explicitly held by some of his contemporaries, not simply an implicit consequence of their arguments. For this reason it seems reasonable to call this thesis an independent thread of Frege's vision of psychologism. It seems unlikely, however, that anyone has ever held a view as simple-minded as the one Frege attacks. In the *Grundgesetze der Arithmetik* Frege foists the equations of truth and acceptance off on Benno Erdmann's *Logik* (1892). He then presents such simple criticisms of the idea that one doubts that there was anyone out there who needed to be told these things. "If it is true that I am writing this in my chamber on the thirteenth of July 1893, while the wind blows out-of-doors," Frege lectures, "then it remains true even if all men should subsequently take it to be false." (1893/1964, xvi).

The final thread in his contemporary's epistemology that Frege objected to was the belief that logic was an empirical discipline that described observable regularities in human thought. According to Frege, if the logician simply describes existing patterns of thought, she will not be able to hold up the canons of logic as rules for proper cognition. The laws of thought a descriptive logician discovers will only be normative in a

derivative sense. Empirically discovered laws might provide norms “in the sense that they give an average, like statements about ‘how it is that good digestion occurs in man’, or ‘how one speaks grammatically’, or ‘how one dresses fashionably’” (1893/1964, xv). The content of these norms is derived from descriptions of human behavior; therefore, their force is not objective, because humans change their behavior constantly. “Just as what is fashionable in dress at the moment will shortly be fashionable no longer and among the Chinese is not fashionable now, so these psychological laws of thought can be laid down only with restrictions on their authority” (ibid.).

The first thing to note about Frege’s notion of psychologism is that its most prominent tenet is not a feature of Kitcher’s psychologism. Kitcher does not say anything about the relationship between meanings and mental entities. This idea stands out so much in Frege’s writings that it is often taken to *be* his notion of psychologism.⁵ Moreover, Kitcher clearly does not make the naïve mistake of confounding truth with acceptance. His analysis of knowledge clearly separates the belief that p from the fact that p is true. Nevertheless, Kitcher does not call his position psychologism without reason. By looking at the way a belief was formed in order to determine if it has the status of knowledge, Kitcher is advocating the use of genetic stories of the sort Frege found objectionable. Kitcher’s psychologism also openly advocates the use of descriptive methods. In doing so, he opens himself up to the charge that he will be unable to develop

⁵ Michael Dummett, for instance, defines psychologism as, “the intrusion or appeal to mental processes in the analysis of sense” (1973, 659).

objective norms. Thus Kitcher seems to conform to Frege's stereotype of the psychologist in two related ways, by advocating genetic accounts and by advocating descriptive accounts. Let's look at these one at a time.

We saw that Frege's specific objections to Husserl's genetic account of mathematics involved Husserl's use of mental entities to stand for mathematical objects, an issue Kitcher avoids. But Frege had a more general objection to developmental accounts of mathematics: they confound the context of justification with the context of discovery. This is an objection Kitcher is ready for. He points out that in his epistemological framework, it is still possible for a belief to be initially held accepted for poor reasons, and only later given an adequate proof (1984, 16–17). These different events can be thought of as a separate "context of justification" and "context of discovery." Indeed, Frege often talks about justification and discovery in a very similar manner, citing the fact that rigorous proofs often only come long after the discovery of a proposition as evidence that discovery and justification are distinct concepts (1884/1968, §3). We can spell out Kitcher's version of this distinction in terms of his analysis of knowledge. Under Kitcher's analysis, a person knows some proposition if the proposition is true, she believes it is true, and that belief was formed by a warranting process. We examine the context of discovery when we consider any process which leads to belief; we examine the context of justification when we consider only warranting processes.

Kitcher's response to the issue of normativity is more detailed. In "The Naturalist Returns" he essentially argues that a psychologistic epistemology will be *better* at

developing a normative theory than an apsychologistic one. The force of the normative claims of the apsychologistic epistemologist is generally taken to come from the notion that she is offering an analysis of the meaning of terms like 'reason' or 'justification'. The sorts of structures of belief she recommends simply define what it is to be rational, and if one wants to be rational, one should structure one's beliefs accordingly. But, Kitcher points out, "Why should we care about these concepts of justification and rationality?" (1992, 63). The psychologistic epistemologist is in a much better position to demonstrate the significance of her normative claims. When she examines a method for acquiring knowledge, she can explain why that method is reliable in a certain environment. This option is not open to the apsychologistic epistemologist. Descriptions of the world and our actual interactions with it play a necessary part in any evaluation of the reliability of an epistemic practice. Often we simply need to know some information about the way the world is, and the way we are, before we know what sorts of tactics will be useful in discovering information about the world, and whether we will be capable of implementing these tactics (*ibid.*, 85). For instance, to determine the optimal composition of a team of researchers, studies have been performed that traced teams of scientists working on identical problems. The normative claims of an epistemologist who had performed such a study will have more weight than the claims of an epistemologist who is merely presenting an analysis of her concept of rationality. In general, historians of science have complained that philosophers take no account of what successful scientists have done or are capable of doing before they make their prescriptions about scientific

method. The psychologistic epistemologist is able to give her prescriptions more weight because she avoids this pitfall.

On some level, this last point about implementing the recommendations of the epistemologist shows that Kitcher and Frege are talking past each other. One does not need to know much cognitive science to see that one cannot balance one's checkbook using the principles of the *Grundgesetze*. (And not just because it's inconsistent!) Frege was not putting forward a normative account in the sense that he was making suggestions for how people should go about their lives. He was identifying the logical principles which underlie our arithmetic knowledge, and his account was normative in the sense that it was the purest expression of rational thought as we understand the term 'rational.' All this, I think, can be conceded by the psychologist, because however Frege understood the normative import of his own work, it is now clear that the psychologistic epistemologist can also produce work of normative value. Thus the charge that psychologism replaces a normative theory with a purely descriptive one is unfounded.

Some of the interplay between norms and descriptions can be seen in Kitcher's use of the history of mathematics in his study. Kitcher's naturalism obviously requires the use of the history of mathematics in the epistemology of mathematics: "To understand the epistemological order of mathematics one must understand the historical order," he says (1984, 5). This does not mean, however, that the norms of mathematics are slave to history. Because he can distinguish between early, inadequate proofs and later, adequate ones, Kitcher can point to places where the historical order diverges from the

epistemological order. On the other hand, since he is in the business of showing the reliability of the processes by which we have developed our mathematical knowledge, he cannot completely disregard the order of history in building the epistemological order. For Kitcher the epistemological order must, “follow the historical order *grosso modo*” (ibid., 9). In fact we shall soon see that Kitcher’s account of mathematics can both link the legitimization of mathematical knowledge to empirical investigations he believes occurred millennia ago, yet still produce a modern axiomatization of arithmetic. Kitcher thus uses Frege’s preferred epistemological method—rational, axiomatic reconstruction—yet does it in the context of an essentially historical account.

Let’s turn now to the other figure I wanted to consider from the antipsychologistic tradition of the nineteenth and twentieth centuries, Edmund Husserl. Although Husserl’s rhetoric mirrors the rhetoric of antipsychologistic writers like Frege, he actually rejects their argument, advancing his own case against psychologism. Husserl begins by assuming that there is a logic that is a practical technology of thought, then attempts to show that this technology needs an a priori, antipsychologistic foundation (1900–1/1970, I 56). His first step in doing this is to point out that any technology has a normative component—it assumes that a certain end is worthy of pursuit and suggests means to that end. Husserl then offers an analysis of normativity according to which any normative claim presupposes a descriptive one (ibid., 81 ff.) Consider a norm like “philosophers should be wise.” A statement like this does more, Husserl claims, than exhort philosophers to be wise, because it can be made without reference to any particular

person's wishes. From this Husserl infers that in making a normative statement we must have an ideal image in mind, such as the ideal philosopher, to which we are comparing the actual world. The problem with psychologism for Husserl is thus not that it bases its norms in descriptions. Husserl's analysis of normativity actually requires that it does this. In fact when Husserl rehearses the standard debates surrounding psychologism, he brings up Frege's objection that a normative science cannot be founded in a descriptive one, and gives off mild surprise that such an argument was ever offered, "Remarkably enough, the opposition [to psychologism] believes it can base a sharp separation between the two disciplines [logic and psychology] on precisely the normative character of logic" (ibid. 92). The reply he offers to this line of argument is reminiscent of the one Kitcher offers in "The Naturalists Return": "Thinking as it should be is merely a special case of thinking as it is" (ibid., 92). Husserl also rejects the antipsychologistic argument we saw Natorp give earlier, the claim that logic cannot be grounded in a science that it was meant to ground. He points out that one need not know logic to be logical, any more than one need study aesthetics to be a good artist. Therefore one can practice psychology before one develops logic—one can even use psychology to develop that logic.

The real problem with psychologism, according to Husserl, is that it relies *solely* on the descriptions provided by empirical psychology. Empirical psychology is inexact and uncertain, and any logic based on it could contain only probabilistic, contingent statements. But logic, for Husserl, is exact and necessary. Therefore, while it might not be completely isolated from psychology, it must have foundations outside of it.

Arguments for psychologism have shown, Husserl admits, that “psychology *helps* the foundation of logic,” but not that “it provides logic’s *essential foundation*” (ibid., 96). Thus the problem is not that psychologistic logic rests on descriptions, as Frege claims, but that it does not rest on *a priori* descriptions of the sort that would come to characterize phenomenology. This difference between Husserl and Frege leads Mohanty to describe Husserl as advocating a “weak psychologism,” even at the time of the *Logical Investigations* (1980, 21).

To the arguments from the modal status and mathematical rigor of logic, Husserl adds two more lines of reasoning. First, he claims that if logical laws were like psychological laws, they would make empirical predictions, which they clearly do not (1900–1/1970, I 104 ff.). Second, he claims that psychologism is a form of relativism, and that all forms of relativism are self-refuting. Psychologism, “conflicts with the self-evident conditions for the possibility of a theory in general” (ibid., 135). The psychologistic logician asserts that logical truths are true because of the composition of the human brain. Therefore, the truth of logical statements, and consequently all other statements, is restricted to humans. But no theorizing is possible if “true” only means “true for humans.” Theoretical statements must be taken as true in a universal sense, if they are taken to be true at all. Thus psychologism undermines the conditions for its own possibility.

Husserl’s arguments are thus different from Frege’s, and pose a separate challenge for Kitcher. Fortunately for Kitcher, replies are quick in coming. The first step in Husserl’s argument was to present a theory of normativity that assumed that because normative

statements do not appear to refer to any particular person's wishes, they must rest on descriptions. But to move from the observation that normative statements do not appear to rest on any particular person's wishes to the claim that they in fact do not is an unjustified leap. Husserl certainly owes an argument to people like A. J. Ayer, who have claimed that this part of the appearance of normative statements is deceiving. This gap in Husserl's argument is not a big problem because he only used the claim that norms rest on descriptions to justify his rejection of Frege's vision of antipsychologism. The real problem with Husserl's arguments is that they rest on assumptions that a modern psychologistic epistemologist need not accept. His first two arguments are built on the idea that logical statements are exact and necessary. However in our wild, post-Quinian, world, it is perfectly respectable to say that logical statements are not qualitatively different in necessity or exactness than any other kind of statement. Kitcher also need not assume that logic makes no empirical predictions. Indeed, Kitcher assumes that logic makes a very specific kind of empirical prediction. Logic for Kitcher is constantly making statements of the form, "If one thinks like this one will achieve one's epistemic ends." What varies is the method of thought and the epistemic ends referred to. Finally, Kitcher need not accept Husserl's conditions for the possibility of all theorizing. It is true that Kitcher's epistemology is based around human capacities and human epistemic ends, but what more could you require of it? Husserl believes that there is a higher degree of universality implicit in every statement that claims to be true. But what more can a statement claim, other than to be valid for those who are capable of understanding it and

expect the same things out of true statements? And this is just to be valid for creatures with human capacities and human epistemic ends.

A Priori versus Empirical

Although Kitcher wishes to say that mathematical knowledge is empirical, he does not wish to do so by declaring that the idea of a priori knowledge is incoherent, or that no a priori knowledge is possible. He instead uses his psychologistic mode of inquiry to expound a new analysis of the distinction between analytic and empirical, and then uses this analysis to frame his argument. Although he does not say so explicitly, I think Kitcher offers his definition of 'a priori' in order to score two rhetorical points. First, he is conceding something to nonempiricists, namely, that an alternative is possible. He needs to do this, because it may seem that by assuming a psychologistic epistemology, he is begging the question against the mathematical apriorist. One might think of a priori knowledge as knowledge that can be seen as warranted without looking at the performance aspects of justification. The way a belief arises can only be relevant to its status as knowledge if it is an empirical belief. But psychologistic epistemology insists that we always look at the causes of belief to determine if it is knowledge, so there is no room for a priori knowledge. Kitcher, by providing a psychologistic definition of 'a priori', shows that it is possible to talk about a priori knowledge even when one must always consider the performance aspect of justification. Kitcher's second tacit motivation is the desire to show off a strength of his psychologistic approach to knowledge: its

ability to provide nice definitions of traditionally contentious terms. To see whether he is successful on these counts, we need to look at the definition he offers. As we saw earlier, Kitcher uses the term *warrant* to refer to a process by which one acquires a belief and which legitimates that belief as knowledge. Armed with this bit of terminology, Kitcher defines a priori knowledge as follows:

- (1) X knows a priori that p if and only if X knows that p and X's belief that p was produced by a process which is an a priori warrant for it.
- (2) \mapsto is an a priori warrant for X's belief that p if and only if a is a process such that given any life e it is sufficient for p that,
 - (a) some process of the same type could produce in X a belief that p
 - (b) if a process of the same type were to produce in X a belief that p , then it would warrant X in believing that p
 - (c) if a process of the same type were to produce in X a belief that p , then p (1984, 24).

Roughly, one has an *a priori warrant* for a belief if and only if one's counterparts in any possible world have available to them a warrant for that belief of the same type as the warrant one has in the actual world. Further riders are put on this definition that restrict the possible worlds in question to those where one's counterparts have all of the cognitive capacities distinctive of humans and have sufficient experience to contemplate the relevant beliefs.

In his actual analysis, Kitcher uses the phrase "any life" rather than the phrase "counterparts in any possible world," which I used in my recapitulation of Kitcher's analysis. My gloss is true to the general way Kitcher talks, however. He slips from referring to "possible lives" where one has a certain experience to referring to "possible

worlds” (e.g., on page 31). No matter what phrase he uses, Kitcher does not mean to introduce any particular modal thesis; neither possible lives nor possible worlds need be construed realistically. The only notion of possibility at play here is the notion already present in the basic conception of a priori knowledge that has been batted around since Immanuel Kant. Kant captured the basic meaning of the phrase “a priori knowledge” when he emphasized that a priori knowledge was “not knowledge independent of this or that experience, but knowledge absolutely independent of all experience.” (*CPR* B3) This remark is already pregnant with the notion of possible worlds. In order to investigate “all experience” one has to think about all the experience that it is possible for humans to have. The only alternative would be to think of “all experience” as referring to “all actual experience,” a concept that could only be investigated if one knew exactly what every actual human had experienced. The need to understand “all experience” as “all possible experience” is also witnessed by Kant’s own method for investigating the a priori, looking for the necessary presuppositions of all possible experience. So despite Kitcher’s talk of possible worlds and possible lives, he is not introducing any modal theses alien to the notion of apriority.

It might be argued that Kitcher has failed to capture the traditional notion of a priori knowledge in his definition because the phrase ‘a priori knowledge’ will not wind up having anything near the extension it is ordinarily thought to have. Obviously, mathematics will not be in its extension. As we will soon see, linguistic truths are not either. The only candidate Kitcher proposes for a priori knowledge is the proposition “I

exist”: “Although traditional ideas to the effect that self-knowledge is produced by some ‘non-optical inner look’ are clearly inadequate,” Kitcher writes, “I think it is plausible to maintain that there are processes which do warrant us in believing that we exist—processes of reflective thought, for example—and which belong to a general type whose members would be available to us independently of experience” (ibid. 29). This is an odd proposition to be called a priori, if only because it is contingent. Although Kripke has argued for the possibility of a priori contingent knowledge, it is certainly an odd turn of events when the leading candidate for a priori knowledge turns out to be a contingent statement. But it should not really trouble us that once all the arguments are in Kitcher’s a priori does not cover the same territory as other’s a priori. If Kitcher’s analysis of ‘a priori’ directly implied that mathematical and linguistic truths are not a priori, he would be begging the question. But Kitcher’s analysis does not directly imply such things. We still have a lot more argument to go. At this point the extension of ‘a priori knowledge’ is completely open, which is how it should be. Really, the judge for whether Kitcher has captured our intuitions about the a priori is the intension of his definition. Here Kitcher is on stronger ground. His definition is an elaboration on the idea that a priori truths are ones that one knows independently of experience, that is, those truths one would know, whatever experiences one had.

Analytic versus Synthetic

Kitcher's treatment of the analytic-synthetic distinction follows the same pattern as his treatment of a priori knowledge. Rather than rejecting the distinction, as many since Quine have done, he expounds his own psychologistic version of it and uses that version to advance his argument. He is more explicit this time about his reasons for taking this tack. He thinks that those who believed mathematical knowledge to be a priori because it is analytic have not gotten a fair shake. If one adopts a psychologistic vision of analyticity, then many of the standard complaints against the analytic-synthetic distinction fall apart. Kitcher does not believe that mathematics is empirical because he thinks the analytic-synthetic distinction fails. He thinks it is empirical because he thinks analytic knowledge is empirical. There is a slight lacuna in his argument here, however. As I will now argue, although he successfully defends a notion of analyticity, he does not show that all analytic statements are empirical. He merely shows that some are. This will be important later on, because it means he will need to show that mathematical statements are analytic statements of the empirical sort.

An analytic statement is typically thought to be one that is true in virtue of the meaning of its terms. In typical psychologistic fashion, Kitcher shifts the focus away from the terms used and towards the knower who grasps those terms. According to Kitcher analytic knowledge (in his idiom, "conceptual knowledge") is knowledge that is warranted solely through the exercise of one's linguistic capacities. "When we learn a language," Kitcher explains, "a complex set of dispositions is set up in us. In virtue of the

presence of these dispositions, which comprise our linguistic ability, we become able to entertain certain beliefs. Let us now suggest that exercise of our linguistic ability generates in us particular beliefs and that it warrants those beliefs” (ibid., 70). The set of these beliefs closed under logical consequence is the set of analytic truths, where logical consequence itself is defined in terms of the exercise of linguistic ability. Exactly how exercise of our linguistic ability is supposed to generate and warrant beliefs is unclear, however. Kitcher suggests an analogy between the exercise of our linguistic abilities to identify a sentence as analytic and our ability to identify a sentence as syntactically well formed. Just as exercising one’s linguistic abilities on a sentence can lead one to know that the sentence is syntactically well formed, so could exercising those abilities lead one to know that it is analytic (ibid., 71 ff.). The linguistic ability is simply expanded to include semantics. This comparison indicates that Kitcher is thinking about a sort of rational intuition, underpinned by a causal mechanism lying somewhere in the language centers of our brain, which allows us to “see” that a sentence simply must be true because of the meanings of the terms involved.

On the surface, this definition of analyticity seems no more immune to the central Quinian objections than any other. The chief objection Quine leveled against the old versions of the analytic-synthetic distinction was that they could not be defined in a non-question-begging fashion. In order to explain what the difference between analytic and synthetic truths was, one had to employ a prior notion of analytic and synthetic. This was the thrust of the first two thirds of Quine’s “Two Dogmas of Empiricism.” Kitcher’s

distinction seems to fall victim to this argument as thoroughly as the traditional positivist's. He says that analytic statements are those known through the exercise of purely linguistic capacities, but how can purely linguistic capacities be identified? It seems unlikely that we could define purely linguistic capacities in a noncircular fashion, especially given the Quinian thesis that there is no principled way to separate meaning and belief.⁶ Think of the classic Quinian example of attempting to learn the word 'rabbit'. In acquiring the ability to use the word 'rabbit' I will acquire a disposition to say things like, "There is a rabbit over there," when there is a rabbit over there. But suppose that while I am learning to use the word 'rabbit' I also acquire the belief that a certain kind of fly is always accompanied by a rabbit. Then I will also have a disposition to say, "there is a rabbit over there," when I see that kind of fly. To separate the dispositions that constitute the meaning of a term from dispositions that represent collateral beliefs acquired during language acquisition would require a prior notion of what sentences were analytically true. This problem becomes critical for Kitcher when we remember that Kitcher needs to provide an account of analytic *knowledge*. It is not enough to define a class of beliefs as those beliefs generated in a certain way, he needs to say why these beliefs are warranted. To do this, he separate out genuinely linguistic dispositions from the dispositions involved in other beliefs acquired during language acquisition, because these collateral beliefs might well be false. If they are false, and their corresponding dispositions are included in the set of dispositions Kitcher counts as a part of our

⁶ This point was emphasized to me by Michael Williams.

“linguistic capacities,” the exercise of one’s linguistic capacities will look less and less like a warranting process.

Kitcher’s reply to this objection is rather unsatisfactory. He assumes that any beliefs acquired while acquiring language will be beliefs accepted on authority from those teaching one the language. So the issue in his mind is distinguishing knowledge gained through exercise of linguistic capacities and ancestral lore. I am not sure that this is the right way to frame the problem, but will grant it anyway. Once the issue is framed this way, he offers two responses. First, he says that we are able to intuitively tell the difference between beliefs generated by our linguistic capacities and beliefs told to us by others. But while there are some cases where this might be true, there are also plenty of hard cases. I am not sure, for instance, if it is a part of the definition of a black hole that nothing can escape from it, even light, or if this is merely something I have been told. Second, he says that some of our beliefs accepted on authority will be remnants of things stipulated to be true. This second point is something that Quineans acknowledge because they don’t think that it will allow for the range of analytic truths that the proponent of analytic truth wants. Kitcher also thinks it will not buy all the knowledge people want out of analyticity, so he and Quine are in basic agreement here. But if Kitcher’s account of analyticity offers no improvement on Quine, why go through all the trouble?

Kitcher is in a much stronger position to respond to these objections once he presents his argument that all analytic statements are empirical. Kitcher’s motive for saying that all analytic statements are empirical comes less from Quine than it does from Mill. Mill

(1843, 91 ff.) argues that statements true in virtue of the definitions of their terms are known empirically because the definitions one operates with ought to be warranted by experience. Mill illustrates the need for such a warrant using the case of the definition of acid. At one point it was a part of the definition of 'acid' that all acids contained oxygen. The use of this definition was no longer warranted once a chemical (hydrochloric acid) was discovered that had all the properties of acids, save for the fact that it contained no oxygen. It was much less of a disturbance to our system of beliefs to change the definition of 'acid' than it was to invent a new category to describe the freshly discovered oxygenless compound. Thus the definition was changed. Kitcher uses a version of Mill's argument, presented in his own terminology, to show that all analytic statements must be empirical. This argument also has the effect of putting to rest lingering Quinian concerns we might have about distinguishing the exercise of a person's linguistic abilities from any collateral beliefs they might have. Essentially, what Kitcher does is allow that analytic statements can include things that look to someone who does not use the language in question like ordinary belief statements—even false belief statements. Such statements are still true in virtue of the language they are couched in, and known to be true because of our linguistic abilities. In the language of chemistry before the discovery of hydrochloric acid, the claim that all acids contain oxygen is analytically true. The paradox of having a statement that is both analytically true and empirically false is avoided by saying that the language that the statement 'all acids contain oxygen' is true

in is an inadequate language for describing the world, thus the terms used in it are nonreferring.

Kitcher's exact argument that analytic statements are empirically known runs as follows. First, Kitcher distinguishes two kinds of revision of language: strong revisions, where the language changes in such a way that the negation of a previously used sentence is now used; and weak revisions, where a sentence simply drops out of use (1984, 81). Kitcher, arguing on behalf of the proponent of a priori analytic knowledge, claims that strong revisions can never occur for sentences thought to be true in virtue of their linguistic properties. If we regard a sentence in the new language as the negation of an analytic sentence in the old, we are simply mistranslating between the new and old languages. Kitcher concedes this because he believes that weak revisions in analytic statements are enough to dislodge analytic statements from their a priori status. Weak revisions do not just occur when a piece of jargon falls out of fashion. Often changes in our empirical knowledge rationally obligate us to drop a definition—we lose our warrant to talk a certain way. Such was the case, according to Kitcher, with the definition of acid. At first "all acids contain oxygen" was an analytic statement, part of the definition of acid. Later the sentence "not all acids contain oxygen" came to be held true. This change, however, can only be thought of as a strong revision of the language if one assumes that the later word "acid" is an adequate translation of the earlier word. But, because we always preserve logical truth in translation, we cannot equate the two uses of "acid." There are really two terms: acid₁ and acid₂. When the language of chemistry

changed we lost the warrant to use the term acid₁ and gained a warrant to use the term acid₂. Now according to Kitcher, when we lose our warrant for using a language we also lose our warrant for any belief acquired through the use of the linguistic abilities which were behind that language. Therefore we have lost our warrant to believe that all acids contain oxygen, and our knowledge that all acids contain oxygen is shown to be a posteriori. We do not have a warrant available for the statement “all acids contain oxygen” in worlds where we are not warranted in using the old chemical language. In general, the knowledge we gain by exercising our linguistic abilities does not meet Kitcher’s criteria for a priori knowledge. As he puts it, “The warranting power of exercising one’s linguistic abilities can be subverted by lives where it is not rational to use the language in question” (ibid., 82).

This is a radical conclusion, so it would be good to scrutinize each step in Kitcher’s argument. As I see it, there are four basic steps.

1. Linguistic abilities only have warranting power to the extent that the use of the language itself is warranted.
2. The warrant for using a language is empirical.
3. If the warrant for the language is empirical, then the warrant for propositions known through the exercise of that language is also empirical.
4. Therefore analytic statements are empirical.

Justification for the first step of the argument is basically an appeal to the pragmatic function of justification. One thing our warrants must allow us to do is convince others that we are right. But our appeal to the meaning of terms will have little effect if the meanings we use are out of favor: “While appeal to linguistic understanding can serve as

a *local* justification for belief, empirical discoveries are relevant to the continued success of the appeal.” Kitcher’s remaining steps are direct applications of his definition of a priori warrants. If in some possible worlds we have no warrant to use a language, then our warrant for that language in this world is not a priori. Further, if our warrant for analytic beliefs only extends as far as our warrant for using the language that makes those beliefs analytic, then our warrant for those analytic beliefs does not exist in all possible worlds. Therefore, our warrant for believing in analytic statements is not a priori.

It should be noted that this whole argument only makes sense in the context of Kitcher’s psychologism. In an apsychologistic context, one can only talk about the justification of sentences by their relationship to other sentences. The use of a language cannot be justified, nor does it stand in need of justification. Therefore, whatever arguments one might produce to legitimate the use of a language, they will not be in the same category of reasons as empirical justifications. The logical positivists are a prime example of a group of apsychologistic epistemologists who denied that the use of a language required empirical justification. Rudolf Carnap, for instance, admitted that many empirical facts might be weighed when deciding to use a language, but because he believed that only sentences in a logical calculus admitted of true justification, he always insisted that the decision to use a language was qualitatively different from empirical discoveries. This line of thinking can be seen in his discussion of the language of material objects, the so-called thing language, in his essay “Empiricism, Semantics and Ontology”: “The thing language in the customary form works indeed with a high degree

of efficiency for most purposes of everyday life. This is a matter of fact, based upon the content of our experiences. However, it would be wrong to describe this situation by saying: ‘The fact of the efficiency of the thing language is confirming evidence for the reality of the thing world’; we should rather say instead ‘This fact makes it advisable to accept the thing language.’” (1950/1988, 208)

This argument does more than simply show that analytic statements are not a priori. It effectively co-opts the final third of “Two Dogmas of Empiricism” to show that a working conception of analytic knowledge is possible. Earlier it seemed as though the kind of arguments raised in the first two thirds of “Two Dogmas” would overwhelm Kitcher’s understanding of analyticity. Kitcher defined analytic sentences as those known through exercise of our linguistic capacities, but there seemed to be no way to define purely linguistic capacities without first assuming some notion of analyticity. The difficulty lay in trying to separate the dispositions stemming from possibly false collateral beliefs from the dispositions involved with the meaning of a term. Once we allow analytic knowledge to be empirical, however, we can simply bite the bullet and let the dispositions involved in ‘false beliefs’ into our understanding of our linguistic capacities. Thus the dispositions associated with “all acids contain oxygen,” are part of the linguistic capacities of a chemist prior to the discovery of hydrochloric acid. One might think that this would lead to relativism, where truth changes as our language changes, so that at one time it is true that acids contain oxygen and at another time it is not. Kitcher avoids this by saying that the two words “acid” are not the same. It is true of acid₁ that it contains

oxygen—analytically true no less. But, as it turns out, there is no acid₁ in the world. Once we are able to include dispositions that look like the basis for false beliefs in our definition of a person's linguistic capacity, it is not difficult to define a person's linguistic capacities in a noncircular fashion. Roughly, the capacity will consist of those dispositions acquired in the use of a language that are required to communicate with the other members of one's linguistic community. Thus dispositions surrounding "all acids contain oxygen" are a part of the linguistic capacities of a chemist prior to the discovery of hydrochloric acid because they form a part of his ability to communicate with his fellow chemists.

It might seem that in making this move Kitcher has failed to produce anything resembling the traditional analytic-synthetic distinction. After all, isn't Quine's claim that all statements are revisable the essence of his rejection of analyticity? Actually, it isn't. Even for Carnap one can revise analytic sentences by changing languages. Kitcher is not countenancing any ability to revise analytic statements stronger than this. In all other respects Kitcher's definition looks like an elaboration of the definitions of analyticity which have been proffered since the sixth century. In his treatise *How Substances Can Be Good in Virtue of Their Existence Without Being Absolute Goods*, Boethius spoke of "a statement generally accepted as soon as it is made" because "no one who grasps [the proposition] would deny it" (*TT*, df. 1). Similarly, under Kitcher's notion of analyticity, the same capacities that allow you to grasp a proposition also tell you that it is true. It is true that Kitcher's definition of 'analytic' is more inclusive than typical definitions. For

starters, it allows not only statements like ‘all bachelors are unmarried’ into the analytic camp, but also sentences that reflect deeply ingrained beliefs about objects, such as ‘all acids contain oxygen.’ More troublingly, statements surrounding stereotypes associated with terms may end up being analytically true, because one’s ability to communicate using a term often depends on having the same associations with that term as others. It would be odd if ‘bachelors tend to live in dirty apartments and eat leftover take-out food’ turned out to be analytic. The situation gets worse when one considers pernicious and false stereotypes. As it stands though, I think this is a matter for further investigation, rather than a refutation of Kitcher’s notion of analyticity.

Although I believe that Kitcher’s argument shows that a notion of analyticity is possible, I do not think it shows that all analytic statements are empirical, merely that some analytic statements are. There are two ways of seeing how some analytic statements may still be a priori. First, suppose there are features that all languages have in common and that exercising the linguistic abilities necessary to master these features could lead to knowledge of certain propositions. If this were the case, then this knowledge would be a priori. The propositions known in this fashion would not be able to undergo even a weak revision. The terms used in these propositions would be present in every language and could always be combined to make an analytic proposition. Now it is clear that not all propositions known by exercise of one’s linguistic abilities could fall into this category, otherwise all languages would be the same. This means that at least some analytic knowledge is empirical, which is sufficient to deal with the Quinian objections to

analyticity. Nevertheless, there is at least the possibility that some analytic statements are a priori.

The second way of seeing how analytic statements may be a priori is to return to Kitcher's notion of a weak revision of a language. In introducing the notion of a weak revision in a language, I hinted at a distinction between two kinds of weak revisions. On the one hand, a weak revision might occur when a piece of terminology simply drops out of fashion, or the occasion to use it ceases to come up. Here the warrant for using a language does not go away—nothing has made the language illegitimate. It simply happens, by some accident of history, that the language is no longer used. Imagine, for instance, that an elaborate taxonomy was developed for understanding the species of birds on a certain island, but then people stopped visiting the island because it was hard to get to and not all that interesting to begin with. Let's call this sort of change a *very weak revision*. On the other hand, a way of talking might be left behind because experience mandates it. Such was the case with the definition of acid we discussed previously. This is the sort of revision Kitcher has in mind when he talks about "weak revision," so let's retain his original phrase for this sort of change. Now given the wide variation in languages, and the absence of logical vocabulary from many languages, it is unlikely that there are sentences that are immune to both kinds of weak revision. For one reason or another, people end up dropping or never developing words for any concept you care to think of. However, it may be the case that some sentences only undergo very weak revision. It is only by historical accident that such sentences are dropped or fail to

develop. Experience in the world never mandates that one cease to use the concepts involved.

If there are analytic sentences that are only very weakly revisable, then those sentences are known a priori by Kitcher's definition of a priori. For Kitcher, a sentence is a priori if one has a warrant *available* to believe that proposition in any possible world. If an analytic sentence is only very weakly revisable, then one always has such a warrant available. Consider again the taxonomy developed to describe the birds on an island that no one visits or cares about. The taxonomy remains the best way to describe these birds, even if no one cares to use it. Now suppose there are sentences that are analytic in this language—for instance, "All *blue-beaked sap-suckers* suck sap with their blue beaks." A sentence like this will be believed to be true by anyone who is capable of using the taxonomy, and this belief will be brought about by that person's ability to use the taxonomy of the island birds. Now in general a process is a warranting process if and only if it is a reliable route to knowledge. For analytic beliefs, the process that generates them is a warranting process if and only if the language the belief is analytic in is the best one for describe the part of the world it is meant to describe. But this is exactly the case here. So statements that are analytic in the taxonomy of the island birds are known a priori. More generally, statements that are only very weakly revisable are known a priori. Again, this argument does not show that all analytic statements are a priori, because presumably some statements will be weakly revisable. However, it does create an opening for a priori analytic knowledge.

How Statements of Basic Arithmetic and Set Theory Are Known

These definitions of a priori and analytic knowledge set the stage for Kitcher's empiricist account of mathematics. I will begin by outlining his account of basic arithmetic. His account of set theory is more complicated and has problems that would cloud the basic structure of his position. In his account of basic arithmetic, Kitcher can be seen as quietly mending problems with Mill. Mill argued that mathematics is empirical by claiming that its axioms were discovered and legitimated by induction from experience, for instance by generalizing our experiences with manipulating pebbles. Kitcher refines this argument by replacing induction with a process of idealization and stipulation. The obvious problem with the claim that the axioms of mathematics are induced from experience is that finite human beings are not capable of gathering enough pebbles to induce the ideas that lie at the heart of the more abstruse regions of mathematics, or even to verify that the basic principles of arithmetic hold for very large numbers. Kitcher rectifies this by asserting that the statements of basic arithmetic refer to an ideal subject who is free of the accidental limitations placed on humans—limitations such as mortality, limited attention span, etc. The ideal subject is capable of two sorts of activities, collecting objects and correlating them with one another, whose natures are specified in the fifteen axioms of what Kitcher terms "Mill arithmetic" (1984, 113 ff.). Let Ux , Sxy , $Axyz$, Mxy be primitive predicates, read " x is a one operation," " x is the

successor operation of y ,” “ x is an addition on y and z ,” and “ x and y are matchable.” The axioms of Mill arithmetic are then

- M1. $\forall x Mxx.$
- M2. $\forall x \forall y (Mxy \rightarrow Myx).$
- M3. $\forall x \forall y \forall z (Mxy \rightarrow (Myz \rightarrow Mxz)).$
- M4. $\forall x \forall y ((Ux \wedge Mxy) \rightarrow Uy).$
- M5. $\forall x \forall y ((Ux \wedge Uy) \rightarrow Mxy).$
- M6. $\forall x \forall y \forall z \forall w ((Sxy \wedge Szw \wedge Myw) \rightarrow Mxz).$
- M7. $\forall x \forall y \forall z ((Sxy \wedge Mxz) \rightarrow \exists w (Myw \wedge Szw)).$
- M8. $\forall x \forall y \forall z \forall w ((Sxy \wedge Szw \wedge Mxz) \rightarrow Myw).$
- M9. $\forall x \forall y \sim (Ux \wedge Sxy).$
- M10. $(\forall x (Ux \rightarrow _x) \wedge \forall x \forall y ((_y \wedge Sxy) \rightarrow _x)) \rightarrow \forall x _x$
(for all open formulas $_$).
- M11. $\forall x \forall y \forall z \forall w ((Axyz \wedge Uz \wedge Swy) \rightarrow Mxw).$
- M12. $\forall x \forall y \forall z \forall u \forall v \forall w ((Axyz \wedge Szu \wedge Svw \wedge Awyu) \rightarrow Mxv) .$
- M13. $\exists x Ux.$
- M14. $\forall x \exists y Sxy.$
- M15. $\forall x \forall y \exists z Azxy.$

The first seven axioms are designed to implicitly define the primitive predicates, the next five are adaptations of the first-order Peano-Dedekind axioms, and the last three are existence assumptions necessary to rule out finite models. He leaves out multiplication; however, it is easy to add a predicate P and two new axioms,

- M16. $\forall x \forall y \forall z (Uy \wedge Pxyz \rightarrow Mxz).$
- M17. $\forall x \forall y \forall z \forall v \forall w ((Pxyz \wedge Syw \wedge Pvwz) \rightarrow Axvz).$

Because the statements of mathematics refer to an ideal entity defined by these axioms, mathematical propositions are true simply by stipulation—we posit the existence of an ideal entity that fits the nature of the axioms, and this entity guarantees the truth of the axioms and their consequents. This tactic of claiming that the axioms of mathematics are stipulated to be true at first seems like something a conventionalist would say, rather than an empiricist. However, the stipulations of Mill arithmetic are not made arbitrarily, or even pragmatically. They are specifically an attempt to model the behavior of actual subjects. Kitcher, moreover, doesn't just believe that his axioms are stipulated truths designed to approximate the behavior of actual subjects. He thinks that mathematics has always worked this way. His axioms are only different in that they are designed to do so explicitly. Now the claim that all mathematical theories are designed to model actual subjects would be easy to make if, in his discussion of analyticity, he had shown that all analytic statements were empirical. All Kitcher would need to claim is that mathematics has always been based on stipulated axioms, and then it would follow naturally that these axioms are empirical. However, as I have argued, there are a couple of possibilities Kitcher needs to rule out here. An analytic statement may still be a priori if it is based on features present in all languages or is only capable of undergoing very weak revision. For Kitcher to declare that mathematical statements are stipulated to be true, yet empirical, he needs to rule out the possibility that they fall into these categories. I would make this argument for him, but I don't think it's true. I'm pretty sure that mathematical statements can always be added to any language without conflict; hence they are only very weakly

revisable. I will only have the resources to show this, however, in Chapter 5. In the meantime, I will grant that mathematics is weakly revisable. It will not make a difference for my critique of empiricism in the next chapter or defense of it in this one.

Mill arithmetic only covers a small portion of mathematical knowledge. To explain the rest, Kitcher needs to come up with a similar account of set theory. He does this by introducing a whole new axiomatic theory that has Mill arithmetic as a consequence. This new theory postulates an ideal agent that is not only able to collect and correlate objects, but also able to collect and correlate acts of collection and correlation. In order to avoid Russell-type paradoxes, restrictions must be placed on when the ideal agent can collect previous collectings. We can't allow it to collect all collectings that do not collect themselves. Kitcher accomplishes this using an iterative conception of a set, drawn from George Boolos (1971). Acts of collecting are divided into stages, each stage only able to collect the collectings in earlier stages. At the base level only objects can be collected. The actions of the ideal agent are defined using the predicates Cxy , $Pxyz$, Exy , xOs , and sEt , where the variables $x, y, z, w...$ range over sets and objects, and the variable $r, s, t...$ range over stages. The predicates then read “ x is a collecting on y ,” “ x is a pairwise ordering of y and z ,” “ x and y are equivalent,” “ x occurs at stage s ,” and “stage s is earlier than stage t .” Kitcher's axioms governing the succession of stages are simply transliterations, using these predicates, of Boolos's axioms. Let's call the resulting theory Kitcher-Boolos stage theory.

$$A1. \quad \forall s \sim sEs.$$

$$A2. \quad \forall r \forall s \forall t ((rEs \wedge sEt) \rightarrow rEt).$$

$$\text{A3. } \forall s \forall r (sEt \vee s = t \vee tEs).$$

$$\text{A4. } \exists s \forall t (t \neq s \rightarrow sEt).$$

$$\text{A5. } \forall s \exists t (sEt \wedge \forall r (rEt \rightarrow (rEs \vee r = s))).$$

$$\text{A6. } \exists s (\exists t (tEs) \wedge \forall t (tEs \rightarrow \exists r (tEr \wedge rEs))).$$

$$\text{A7. } \forall x \exists s (zFs \wedge \forall t (xFt \rightarrow t = s)).$$

$$\text{A8. } \forall x \forall y \forall s \forall t ((Cxy \wedge xOs \wedge yOt) \rightarrow tEs).$$

$$\text{A9. } \forall x \forall s \forall t (xOs \wedge tEs \rightarrow \exists y \exists r (Cyx \wedge yOr \wedge (t = r \vee tEr))).$$

$$\text{A10. } \forall x \exists y \forall x (Cxy \leftrightarrow (\Box \wedge \exists t (tEs \wedge xOt)))$$

(where \Box is any open formula in which no occurrence of y is free).

$$\text{A11. } \forall s \forall t (tEs \rightarrow \forall x (xOt \rightarrow \Box) \rightarrow \forall x (xOs \rightarrow \Box]) \rightarrow \forall s \forall x (xOs \rightarrow \Box)$$

(where \Box is any open formula that contains no occurrences of t , and $\Box]$ is just like \Box , except for having a free occurrence of t wherever \Box has a free occurrence of s).

$$\text{A12. } \forall x \exists t (xOt) \rightarrow \forall z \exists s (Czw \rightarrow \exists r (wOr \wedge sEr)).$$

Axioms 1–5 establish the order of the stages, saying that no stage is earlier than itself, that the relationship *earlier than* is transitive and connected, that there is an earliest stage, and that each stage has an immediate successor. Axiom 6 asserts that there is a stage that is not the first stage, but is not immediately after any other stage, a stage Δ , or stage “at infinity.” Axioms 7–9 describe the stage at which sets may be formed. The set of axioms named in 10 guarantees that sets are formed at the earliest stage possible, and the set of axioms named in 11 captures the principle of induction. Axiom 12 says that if every set occurs at some stage or another, then for any set z , there is a stage s which occurs after the stages at which z 's components occur. This axiom is actually not included by Boolos in his official list of axioms, because he does not feel it is a necessary part of the iterative

conception of a set. He does, however, state the axiom informally, and he asserts that it is necessary to deduce the axioms of replacement.

Boolos gives derivations of the axioms of ZF set theory using his versions of the 11 axioms given above. To help himself to these derivations, Kitcher need only transliterate them into his preferred symbolism. The resulting axioms of ZF, now considered theorems of Kitcher-Boolos stage theory, are written as follows

The Axiom of the Null Set

$$T1. \quad \exists y \forall x \sim Cyx.$$

The Axiom of Pairs

$$T2. \quad \forall z \forall w \exists y \forall x (Cyx \leftrightarrow (x = z \vee x = w)).$$

The Axiom of Unions

$$T3. \quad \forall z \exists y \forall x (Cyx \leftrightarrow \exists w (Cwx \wedge Cz w)).$$

The Power Set Axiom

$$T4. \quad \forall z \exists y \forall x (Cyx \leftrightarrow \forall w (Cw x \rightarrow Cz w)).$$

The Axiom of Infinity

$$T5. \quad \exists y (\exists x (Cyx \wedge \forall z \sim Cxz) \wedge \forall x (Cyx \rightarrow \exists z (Cyz \wedge \forall w (Cz w \leftrightarrow (Cw x \vee w = x))))).$$

The Axioms of Separation

$$T6. \quad \forall z \exists y \forall x (Cyx \leftrightarrow (Cz x \wedge _))$$

(where $_$ is a formula in which no occurrence of y is free).

The Axioms of Regularity

$$T7. \quad \exists x _ \rightarrow \exists x (_ \wedge \forall y (Cxy \rightarrow \sim _y))$$

(where $_$ does not contain y , and $_y$ is $_$ with y substituted for all free occurrences of x).

The Axioms of Replacement

$$T8. \quad \exists x \forall y (Cxy \leftrightarrow _y) \rightarrow \exists z \forall y (Czy \leftrightarrow _z)$$

(where $_y$ is a sentence and $_z$ is the same sentence with z substituted for y).

To complete the derivation of Mill arithmetic, Kitcher introduces definitions of the predicates of Mill arithmetic using the language of his stage theory.

$$\text{A13. } \forall x(\text{U}x \leftrightarrow \exists y \forall z(\text{C}xz \leftrightarrow z = y)).$$

$$\text{A14. } \forall x \forall y(\text{S}xy \leftrightarrow (\forall z(\text{C}yz \rightarrow \text{C}xz) \wedge \exists u \forall v((\text{C}xv \wedge \sim \text{C}yv) \leftrightarrow v = u))).$$

$$\text{A15. } \forall x \forall y \forall z(\text{A}xyz \leftrightarrow (\sim \exists w(\text{C}yw \wedge \text{C}zw) \wedge \forall u(\text{C}xu \leftrightarrow (\text{C}yu \vee \text{C}zu)))).$$

$$\begin{aligned} \text{A16. } \forall x \forall y(\text{M}xy \leftrightarrow & \exists z(\forall w(\text{C}zw \leftrightarrow \exists y \exists v(\text{P}wuv \wedge \text{C}xu \wedge \text{C}yv)) \\ & \wedge \forall u(\text{C}xu \rightarrow \exists w \exists v \forall s \forall t(\text{C}zs \wedge \text{P}sut \leftrightarrow (s = w \wedge t = v))) \\ & \wedge \forall v(\text{C}yv \rightarrow \exists w \exists u \forall s \forall t((\text{C}zs \wedge \text{P}stv) \leftrightarrow (s = w \wedge t = u)))). \end{aligned}$$

As before, $\text{U}x$ is to be read “ x is a one collecting,” $\text{S}xy$ is to be read, “ x is the successor operation on y ,” $\text{A}xyz$, “ x is an addition on y and z ,” and $\text{M}xy$, “ x and y are matchable.” Given these definitions, it is easy to show that the axioms of Mill arithmetic follow.

The system works, but at the cost of some heavy assumptions. In order to allow for the existence of inaccessible cardinals, the ideal agent must be allowed to perform elaborate transfinite acts of collecting (Kitcher 1984, 146). The ideal agent must also be able to refer to future collectings (even though she is not able to collect them) in order to allow for impredicative definition (ibid., 145). Postulating such abilities draws the ideal agent farther and farther away from any real agent, a fact that was remarked on negatively by many reviewers (Parsons 1986, Maddy 1985, Gillies 1985). The discomfort with the powers of the ideal agent becomes most acute in light of an objection from Charles Parsons. He notes that if we assume that the ideal agent performs one act of collecting at one moment in time, and that time is a continuum of the same cardinality as the real number line, then there are not enough moments in time for the ideal agent to

collect the higher cardinals. Kitcher notes this objection in his book, and can do nothing but concede it. To ameliorate its impact he suggests that the ideal agent lives in a hypertime, similar to our time, save for being highly superdenumerable (1984, 145). Really, though, this only highlights the unnaturalness of the ideal agent.

Adopting an alternative to Kitcher-Boolos stage theory might rid us of all these large assumptions, but the alternative theories are hard to interpret in terms of the ideal agent. Boolos laid out his iterative conception of the set to capture ZF set theory. There are alternatives to ZF, such as Quine's NF and ML (1963). In Kitcher-Boolos theory, as in most forms of ZF, we quantify directly over stages using the set of variables s, t, \dots . Quine's systems do not quantify over stages. Russell paradoxes are avoided by placing strictures on the way certain terms in the axioms of NF and ML are understood. The axioms include terms of the form Fx , which are supposed to represent any formula containing the variable x . Russell paradoxes are avoided by saying that when formulas are substituted in for Fx , indices must be supplied in such a way that the symbol " \in " must always be flanked by consecutive ascending indices. The indices take the place of stages, and the requirement that they be ascending ensures that no set can be formed on sets of the same index. Because Quine's systems never quantify over stages, stages do not exist by his famous criterion for existence. However, from Kitcher's perspective the systems seem incredibly ad hoc. There is no way to explain the indices in terms of the ideal agent. They can only be motivated as a way to simultaneously avoid Russell

paradoxes and positing hypertime. Therefore they cannot be seen as a viable alternative to Kitcher-Boolos stage theory.

In general, the choice of a set theory poses a dilemma for Kitcher. Stage theories can be motivated in terms of a ideal agent but grant that agent unlikely transfinite abilities. Stageless theories avoid this difficulty but are impossible to motivate. This dilemma should not come as a surprise: it is an inevitable result of attempting to develop a nonfinitist theory of a mathematical subject. If one is serious about expressing mathematics in terms of an individual subject, one will have to attribute infinite powers to that subject. Because Kitcher is committed to expressing his theory in terms of a subject, his choice in this situation is clear. He must endorse an ideal subject with transfinite powers.

Ontology

Kitcher adds to his epistemology an ontology that gives mathematical objects an imminent, distinctively Aristotelian character. Kitcher views mathematical objects as our way of representing the fact that activities of collecting and correlating yield the results that they do. As he puts it, arithmetic is about “permanent possibilities of manipulation,” or “structural features of the world in virtue of which we are able to segregate and recombine objects” (1984, 108). These structural features are not to be thought of as existing over and above the individual instances of our ability to manipulate objects. Thus Kitcher remarks that he wishes to “replace notions of abstract mathematical objects,

notions like that of a *collection*, with the notion of a kind of mathematical activity, *collecting*” (ibid., 110). On this account, “There is no suggestion of a gap between these ordinary objects [the objects that we manipulate] and other, more ethereal entities which lurk behind them” (ibid., 107).

Kitcher is openly indebted to Mill for this ontology. “Every arithmetic proposition,” Mill wrote, “is a statement of one of the modes of formation of a given number, it affirms that a certain aggregate might have been formed by putting together certain other aggregates, or by withdrawing certain portions of some aggregate, and that by consequence, we might reproduce those aggregates from it by reversing the process” (1843, 430). But to rely on Mill for one’s ontology of mathematics is to ask for trouble. Ontology was the aspect of Mill’s doctrine that earned him the most searing of Frege’s rebukes, and any attempt to rehabilitate Mill’s ontology will have to confront these criticisms. In some cases this will not be difficult. Frege was far from a sympathetic reader, and many of his arguments were anticipated and countered by Mill himself. For instance, Frege took Mill to be claiming that the objects which numerals refer to must be physically collected, prompting him to ask, “Must we literally hold a rally of all the blind in Germany before we can attach any sense to the phrase ‘the number of blind in Germany?’” (1884/1968 §23). As Kitcher has pointed out, however, Mill often talked of numerals referring to aggregates that were only collected in thought, as when he said, “We need only conceive a thing divided into four equal parts (and all things may be conceived as so divided), to be able to predicate of it every property of the number four”

(1843, 189). Similarly, Frege claimed that Mill failed to account for our ability to count intangibles such as proofs of a theorem or events (1884/1968, §23), but if Mill is willing to countenance mentally collecting physical entities, surely he would countenance mental collecting of things that cannot be touched.

Other arguments Frege puts forward are dealt with by Kitcher's use of idealization and the formalism that comes with it. Frege felt that Mill's doctrine lead to absurdities when it was applied to very large numbers: "If the definition of each individual number did really assert a special physical fact, then we should never be able to sufficiently admire, for his knowledge of nature, a man who calculates with four digit figures" (1884/1968 §7). The smallest numbers were thought to pose a problem for empiricism, too. If numbers are properties of physical collections, what collections do the numbers 0 and 1 attach themselves to? (ibid.). Both of these problems are solved by Kitcher's use of idealization. The problem of very large numbers is obviously solved by the use of an ideal agent. However, by defining the ideal agent using formal axioms, Kitcher takes care of the problem of the numbers 0 and 1 as well. Frege's claim that Mill cannot provide an account of the numbers 0 and 1 receives its force from the connotation of plurality carried by the words Mill uses to describe the referents of mathematical language, words like "agglomeration" and "aggregate." Kitcher, however, by outlining his theory in a formal language with primitive predicates, gives himself license to stretch the ordinary meanings of words. Thus Kitcher can introduce the predicate U, indicating an act of collecting one object, without thinking twice. If anyone objects that the phrase "the act of collecting one

object” makes no sense, or is oxymoronic, Kitcher can reply that it is only a heuristic, and that the precise meaning of the predicate “U” is given in the axioms of Mill arithmetic. Besides, formalization always changes the meaning of words a little. Kitcher does not define a predicate that indicates a collection of no objects, but with the appropriate revision of his axioms, he could do so easily. The fact that in ordinary language it makes no sense to talk about collecting zero objects (is it something I do whenever I am not collecting more than zero objects?) will not detract from the rigor of the formal language.

But I am still only responding to Frege’s flippant objections. We must turn to meatier points. Frege’s most famous objection to empiricism, which he learned from Bauman (1868), begins with the observation that no physical object or objects in themselves determine a number. If I place a pile of cards in someone’s hands and ask him to find their number, “this does not tell him whether I wish to know the number of cards, or complete decks of cards, or even say the number of honor cards at skat” (1884/1969, §22). To assign the right number to the pile of cards, one must provide additional information besides the cards themselves. But as Frege points out, “If I can call the same object red and green with equal right, it is a sure sign that the object named is not really what has the green color” (ibid.). Here Kitcher’s Aristotelian ontology serves him well. By claiming that arithmetic is about acts of *collecting* rather than *collections*, he includes the subject doing the collecting in the objects he takes arithmetic to be about. He is asking us to imagine an entity gathering objects, and to do this we must assume that the entity is individuating these objects in some specific way. Thus all of the information

necessary to determine what exactly is being collected is included in his notion of mathematical objects.

Frege also argues that numbers could not be properties of collections in the world on the grounds that grammatically, a number word is a proper name, not a property word. Frege notes that number words do not take the plural, “there are not diverse numbers one, but only one. In 1 we have a proper name, which as such does not admit of a plural any more than ‘Frederick the Great’ or ‘the chemical element gold’” (ibid. §38). Number words also take the definite article, as in phrases like “the number 55” or “the number of moons of Jupiter” (ibid. §§52, 57). We also can talk of numbers being identical to one another, as when we say that “ $2 + 2 = 4$,” or “the number of moons of Mars is two.” This could only be possible, according to Frege, if numbers were objects, and the sentences that asserted identity said that two names picked out the same number (ibid.).

Contemporary linguists (e.g., Hurford 1975) tend to regard numerals in most languages as a unique class of terms, having properties of both adjectives and nouns. The arguments for this position are quite strong. The points Frege brings out make it clear that numerals in German and English often act like proper nouns. However, there are also circumstances in German and English that break this pattern. Sometimes we can form the plural of numbers, for instance when we say “revelers left the party in twos and threes” (Huddleston 1984, 328) Also in many languages small numbers are declined as if they were adjectives (Menninger 1958, 18 ff.). It is doubtful that Frege could discount these

facts without relying on a question-begging distinction between the “real” and “surface” grammar of number words.

We do not need to determine the correct classification for numerals here, however. The question we need to address is whether Kitcher is required to claim that numerals are property words in order to assert that mathematical truths are about “structural features of the world in virtue of which we are able to segregate and recombine objects.” I can’t see why he would be. When Mill formulated his ontology, he claimed that individual numbers referred to collections of objects. Specifically he said that numbers denote physical collections and connote the properties that those collections have in common (1843, 430). Kitcher does not state his ontology in such a simple-minded way. He provides an ontology for mathematical thought in general, without claiming a direct semantic link between individual numbers and physical entities. So, unlike Mill, he does not need to assert that numerals are words for properties of objects. Nor does it seem particularly incumbent on him to explain the relation between the grammatical structure of numeral and the “structural features of the world in virtue of which we are able to segregate and recombine objects.” The relationship will clearly be complex, if not inscrutable, but this is a problem for a general theory of semantics, not for Kitcher’s theory of mathematics.

History

The most intriguing, and difficult to evaluate, aspect of Kitcher's theory is his historicism. In developing his account of mathematics, Kitcher points out that most of our mathematical knowledge is accepted on authority, and that much of the knowledge of our current authorities was passed on from preceding authorities. Under Kitcher's reconstruction of the history of mathematics, the empirical input into mathematical knowledge only occurs at the very beginning. All subsequent changes in mathematical knowledge stem from intratheoretic tension. Unfortunately, Kitcher is unwilling to fill in the details of the empirical origin of mathematics. In *The Nature of Mathematical Knowledge* he suggests, with uncharacteristic vagueness, that the discovery of arithmetic occurred, "several millennia ago, probably somewhere in Mesopotamia" (1984, 5). More recently, he has suggested that the first arithmetic was probably much more ancient than the Babylonian culture he originally alluded to: "It seems to me to be possible that the roots of primitive mathematical knowledge may lie so deep in prehistory that our first mathematical knowledge may be coeval with our first propositional knowledge of any kind" (1988, 322, n. 10). Because arithmetic is so old, Kitcher feels that we are not in a position to say anything about its origins besides the loose remarks he has already made: "although a naturalist contends that mathematical knowledge originated in some kind of response to the environment, it is eminently reasonable to propose that there are a number of possibilities and that this aspect of the naturalistic theory of knowledge is (for the moment, and perhaps permanently) less accessible to elaboration" (ibid.).

Kitcher's defeatism on this issue is probably premature. There is actually a host of evidence that is relevant to the origin of mathematics. If our goal in this chapter is to make Kitcher's views plausible, it is incumbent on us to take at least a brief look at this data. The concept of number grew up independently in Mesopotamia, India, China, and Mesoamerica. Because I can afford to spend so little time on this subject, I will only look at the history of number in Mesopotamia, and look only at one, much debated, reconstruction of that history, and I will still not be able to do it justice. All I hope to do here is give a sketch that will begin to indicate how Kitcher's views on the origin of mathematics is plausible.

The earliest signs of protomathematical knowledge are remnants of people exercising their ability to put sets in one-to-one correspondence. The fossil record shows that people began to make records of one-to-one correspondences on imperishable media in the Upper Paleolithic era, around the same time that they began to paint animals on cave walls and carve the famous "Venus" figurines. These records take the form of thousands of stone tools, cave walls, and fossilized bones which have been marked with various arrangements of notches (Marshack 1972). While there is much debate over the meaning of these artifacts, it is clear that at least some of them are tallies, such as a wolf bone found in the Czech republic dating from 30,000 B.C.E. that was marked with 55 notches, the first 25 of which are arranged in groups of five, with a longer mark after the 25th mark (Sarton 1938; Flegg 1983, 41–42). These most ancient relics do not indicate that the people of the Upper Paleolithic had mathematical knowledge in Kitcher's sense,

however. All that can be said for sure is that the people who carved these bones could place objects in a one-to-one correspondence with a set of marks. For Kitcher, the ability to correlate objects in this fashion is the object of mathematical knowledge, it is not mathematical knowledge itself. To qualify as having mathematical knowledge one must somehow represent to oneself information about what sorts of collections are possible. The notched artifacts of the Upper Paleolithic show no sign of the existence of such a representation. Furthermore, it is unlikely that the people of the Upper Paleolithic had any need of such representations. Humanity in those days lived by gathering and hunting. Contemporary hunter-gatherer societies flourish without any explicit representation of numbers or number properties in their language or systems of abstract thought.⁷ Such cultures do, however, find it useful to employ systems of one-to-one correspondence, such as tally sticks or body counting (for an example, see Saxe 1982). So unless the mechanics of hunting and gathering were different in the Upper Paleolithic than they are today, it is unlikely that the people of the Upper Paleolithic did anything with number besides keep tally sticks. If this hypothesis is correct, then Kitcher is wrong to speculate that mathematical knowledge is coeval with propositional thought. True mathematics is at least younger than the Upper Paleolithic.

⁷ Societies which lack indigenous arithmetic include some native Australian peoples (Dixon 1980, 107–8; Wurm 1972, 63–64; Yallop 1982, 145), the recently contacted people of Papua New Guinea (Saxe 1982, Biersack 1982), and some groups of the San of the Kalahari (Bleek 1928/1978, 38–9; Bleek 1937, 196; Maingard 1963, 102). Of course, one has to be careful here of the long tradition of European anthropologists

The evolution from simply making correspondences to actually understanding them seems to have occurred gradually and moved along predominately through successive realizations that more and more sophisticated sorts of correspondences were possible. The key piece of evidence here is a class of artifacts known as tokens.⁸ Around 8000 B.C.E., during the Neolithic era and the development of agriculture, people in the Near East began to make small clay geometric objects, such as cones, spheres, and cubes. Between 3700 and 2500 B.C.E., when isolated farming settlements grew into towns, tokens were stored in small clay spheres, or bullae. Based on the inscriptions found on bullae made after the invention of writing, it has been established that large numbers of the recovered tokens were a part of an accounting system. In the early stages of the system, individual tokens symbolized individual items, just as the notches on tally sticks did. However it was quickly discovered that tokens of different shapes could be used to signify larger collections. The variety of token shapes (and in the later tokens, patterns of engraving) suggests that the tokens also bore information of the kind of object collected. The bullae were introduced to keep tokens from being tampered with during important transactions. Breaking a bulla was equivalent to breaking a contract. By Kitcher's standards his stage of the development of number is midway between the simple use of one-to-one correspondences and true knowledge of arithmetic. The people of this era

misrepresenting non-European cultures to fit their stereotypes. However, the sources I am using are recent and reliable.

⁸ The information in this paragraph and the following two was gleaned primarily from Nissen et al. (1993) chapters 4 and 16, Nissen (1988), Damerow (1988), and Schmandt-Besserat (1992).

have enough knowledge of the nature of collecting and matching objects to be able to establish values for tokens—ten one-sheep tokens could be equivalent to a ten-sheep token. However, as Peter Damerow has argued (Nissen et al. 1993, chapter 14; Damerow 1988, 138–48), the tokens of this era are not a part of a linguistic system that is capable of representing numbers as abstract entities. Five sheep tokens do not stand for the number five, but for five specific objects. Now, obviously we don't know how the advances of this middle stage in the development of number occurred. However, in general they seem like the sort of steps forward Kitcher's theory would predict. After manipulating collections of tokens for centuries, people began to notice other possibilities of collection, such as establishing different values for different kinds of tokens. It certainly looks like an advance born out of experience with the world.

True numbers arrived in Mesopotamia suddenly around 3100 B.C.E. About that time, climatic changes caused the alluvial plain of the Tigris and Euphrates rivers to recede, drying out the swamps and lagoons between the rivers and revealing the most fertile land the region had seen. The communities around the Babylonian flood plain were already highly urbanized, and when the new land was colonized, this social structure was imported to the new territory wholesale. The city of Uruk, a metropolis over half the size of Rome in the first century C.E., was established as an imperial capitol. The demands on the accountants and bureaucrats of the ruling priestly class were unprecedented. The workload led to rapid changes in the way information about collections of objects was stored and manipulated. Collections of objects such as taxes or disbursements of goods

were symbolized using round and oblong marks in clay tablets. At first, there was no fixed convention for the number value of these marks. But soon other marks began to appear on the tablets—ideograms describing the objects being counted and the nature of the transaction. From the very start, these ideograms were combined systematically to form other ideograms. In essence, writing appeared, and appeared like Athena from the head of Zeus—suddenly and fully formed. Along with the invention of writing came changes in the marks used to represent collections. Number signs acquired a level of autonomy that they did not have before. Conventions for the use of different signs to represent specific values were established. They were given a fixed graphic form and made to stand in a fixed relation to one another. These number signs could be combined with other ideograms to represent, for instance, an animal of a certain age. Most importantly, arithmetic operations appear. If several collections of objects were represented on one side of a tablet, their sum was represented on the other. This was true counting. In Kitcher's scheme of things, the conventions governing the use of number signs play the role of axioms establishing the nature of an ideal accountant. (In Sumarian, an accountant was called a SANGA, .) The theory of an ideal accountant established by these conventions was, like many advances in empirical science, the product of a need to solve a physical problem—how to deal with the masses of goods brought in and out of the city each day. The solution to this problem was discovered by people who worked with the physical objects involved—both the goods and the clay representations of the goods—and were in a position to notice useful properties about

them. Counting was thus an empirical theory, a way of representing the physical world developed by technicians hard at work within that world.

Or so Kitcher would say. In the next chapter I will argue that the empiricist view of mathematical knowledge cannot account for important aspects of mathematical practice. In doing so I will return to Uruk to show that Kitcher cannot use the apparently empirical nature of the discoveries there to support the empirical status of mathematics.

Critique of Empiricism⁹

Introduction

The most striking arguments against empiricism in the twentieth century have come from Ludwig Wittgenstein. While most other thinkers accepted Gottlob Frege's dismissal of mathematical empiricism and thought no more about it, Wittgenstein carefully considered the possibility that we could gain mathematical knowledge through experience. To his mind, the fact that we cannot say something very important about mathematics and the structure of knowledge in general. Mathematical statements lie on one side of a great divide in language: the split between descriptive statements and expressive statements. For Wittgenstein one of the most common causes of philosophical confusion is the mistaken belief that a statement describes an element of the world when it is actually used to express something. For instance, confused beliefs about qualia arise because we think that pain talk describes an entity, pain, when it is really of a piece with expressive behavior like yelps and cries. Mathematical talk, too, is expressive. It belongs to an important class of statements Wittgenstein often calls "grammatical," which are used to express norms, or rules the speaker thinks ought to be followed. Technically, there are two theses at work in Wittgenstein's critique of empiricism, one negative and one positive. First, mathematical statements are not used to describe of the world.

⁹ Portions of this chapter were given as a talk entitled "Dissolving the Application Problem" at the Texas Tech Philosophy Department. I am grateful to the audience at Texas Tech, particularly Aaron Meskin and Howard Curzer, for their helpful criticisms.

Second, mathematical statements are used to express norms. These two claims are often intertwined in Wittgenstein's writing, but in what follows I will try to keep them separate.

Since part of the critique of empiricism is the thesis that mathematical statements are expressions of norms, it is important at the outset to be clear on the terms "expression of norms" and "mathematical statement." There are many ways a statement might be considered an expression of a norm. One might call a statement normative because it describes the norms followed by a community. "India is governed by a complex caste system," is a statement of a norm in this sense. Such a statement is a *description of a regularity* of behavior. A statement might also be an expression of norms because it *entails norms*. "It is raining" might entail "I should take an umbrella." Neither descriptions of regularities or statements that entail norms are expressions of norms in the sense I am using in this dissertation. For my purposes, a statement is an expression of a norm if it *expresses the norm the speaker believes she or someone else ought to follow*. Promises or New Year's resolutions are normative in this last sense because they are means by which the speaker commits herself to an action. A statement is also normative in this sense if it acknowledges a commitment that already exists. If I say, "one ought to be quiet during the movie," I am expressing a norm, but I am not undertaking a new normative commitment. I am acknowledging an obligation that I believe binds all moviegoers. When Wittgenstein calls a statement "normative" or "grammatical," he means that it is an expression of a norm that the speaker thinks ought to be followed, rather than a description of a regularity in human behavior or a statement that entails norms. In most cases, the norm expressed exists before it is expressed. In this sense,

Wittgenstein's grammatical statements are not like promises or New Year's resolutions, which create commitments. Wittgenstein believed that mathematical empiricism was mistaken because it asserted that mathematical statements were descriptions of the world, rather than expressions of norms. He was not always clear, however, whether mathematical statements expressed pre-existing norms or whether they were more like promises and New Year's resolutions, and marked the creation of new norms. In what follows, I will argue that mathematical statements are best viewed as acknowledgments of already existing norms, like Wittgenstein's other grammatical statements.

There is a weak sense in which any assertion is an expressions of a norm. If I say "grass is green" I am committing myself to the belief that grass is green and urging that commitment on others. I am, in effect, saying "We ought to believe that grass is green." Expressions of norms properly so called should urge something more on us than this, however. Any theory of language ought to be able to distinguish between sentences like "grass is green" and "grass ought to be green, even though my lawn is persistently brown." If " $2 + 2 = 4$ " is a normative statement in any robust sense of the term, it must be more like "grass ought to be green." Part of the difference between statements that are robustly expressions of norms and ordinary descriptive statements will clearly lie in how the actual state of the subject of the sentence, the grass or the two collections of two objects, affects our evaluation of the sentence. Exactly what features distinguish normative and descriptive statements will become clear later on in this chapter, when we look at the evidence that mathematical statements are expressions of norms.

Just as it is important to be clear about what we mean by ‘expression of a norm’, it is important to be clear on what we mean by ‘mathematical statement’. When I say that mathematical statements are expressions of norms, I am not just referring to any statement with numbers in it. ‘Mars has two moons’ is a straightforwardly descriptive statement. Roughly, the statements I am concerned with are those of *pure* mathematics, rather than applied mathematics. I acknowledge that this distinction can be difficult to pin down. Statements like ‘The number of Martian moons is two’ look like assertions about numbers, rather than physical objects, but we still do not want them to be parts of pure mathematics. I will not try to come up with an airtight distinction here. Instead I will rely on the rough idea that pure mathematical statements do not essentially refer to some aspect of the physical world. Hopefully this will be definition enough for my purposes.

I believe that the theses “mathematical statements are not descriptions of the world” and “mathematical statements are expressions of norms,” are, in fact, theses, despite Wittgenstein’s famous protestation that philosophy should advance no theses. My general approach to Wittgenstein’s work will be to look at what he does, not what he says he does or what he thinks he does.¹⁰ I think that it is clear that he has offered a thesis here, and an interesting one. I also think that the remarks surrounding this thesis constitute an argument in the ordinary sense of the word: a series of connected statements designed to convince an audience of a proposition. The statements Wittgenstein strings together are

¹⁰ Einstein (1934/1982, 270) said something very similar about the way to understand the scientific method: “If you want to find out anything from the theoretical physicists about the methods they use, I advise you to stick closely to one principle: Don’t listen to their

generally thought experiments. The purpose of these experiments is, Wittgenstein tells us, to elicit our intuitions about the way mathematical language is used. On this matter, Wittgenstein's self-description is accurate. However, the reader must do more than passively note aspects of the use of language. She is supposed to change her beliefs about language in light of what she learns about use. It is in this ability to change beliefs that the argument lies. Wittgenstein tended to leave this key level of his writing implicit—forcing the work of deciding how to change one's beliefs on the reader. In rendering it explicit I will use language alien to Wittgenstein whenever I think it will make his claims more persuasive.

Wittgenstein reached the conclusion that mathematics was normative and not descriptive comparatively early in his career. In the *Philosophical Remarks*, written between 1929 and 1930, he was already making the claim that encapsulates his antiempiricist stance towards mathematics: "The proposition 'vertically opposite angles are equal' means that if they turn out to be different when they are measured, I shall declare the measurement to be in error" (*PR* §178). If we viewed mathematical statements as descriptions, we might think that 'vertically opposite angles are equal' is a kind of prediction; it says what would happen if one measured vertically opposite angles. Here Wittgenstein says that this proposition plays a very different role. He views it as a pledge, which says that if we measured vertically opposite angles and found they had different sizes, we would search for the point where we went wrong in our measurement.

words, fix your attention on their deeds." Thanks to Arthur Fine for pointing out this connection.

But in 1929 Wittgenstein had quite different reasons for holding this view of mathematics than he had just a few years later. Wittgenstein's first argument against empiricism was based entirely in his understanding of the relationship between a theorem and its proof. This argument grew up in the middle period of his life, about the time of the *Philosophical Remarks* and the *Philosophical Grammar*. These themes persisted through the writing of the *Philosophical Investigations* and the notes that were collected in *Remarks on the Foundations of Mathematics*. However, the dominant argument against empiricism in this later period was a series of observations that sprung out of Wittgenstein's famous ideas on rule following. This new argument against empiricism received its most refined statement in the typescript that has since been published as the first book of the *Remarks on the Foundations of Mathematics*. The difference between these arguments is not always appreciated by Wittgenstein's critics, who assume that the flaws of Wittgenstein's argument in 1929 also mar his later argument. In what follows I will run through both the argument against empiricism found in the *Philosophical Remarks* and *Philosophical Grammar* and the argument found in the *Remarks on the Foundation of Mathematics* in order to clearly distinguish the two. To get a clear picture of what I take to be Wittgenstein's better argument, I will restrict my focus to the first book of the *RFM* and only touch briefly on the remarks on rule following in the *Philosophical Investigations* that were meant to lead into Book 1 of the *RFM*. I will also look at some objections to Wittgenstein's arguments. I will not look at all of them, however. In this chapter I will only look at the sorts of objections that might be raised by the mathematical empiricist. More fundamental objections, especially the charge that

Wittgenstein makes mathematics relative, or worse yet, some species of rhetoric, will be put off until the last chapter.

The Middle Period Argument

Wittgenstein's middle period argument against mathematical empiricism is primarily an argument for the thesis that mathematical statements are not descriptive. While he already holds the second thesis, that mathematical statements are expressions of norms, in his middle period, it is not yet developed or argued for. The middle period argument can be neatly summarized like this: the objects of empirical discovery can be characterized apart from their discovery, but a mathematical proposition is inseparable from its proof. As a result, whenever we have an adequate characterization of a mathematical proposition, we know that it is true. This sharp contrast between mathematical and empirical statements keeps mathematical statements from being descriptions. Variations on this argument are presented in the *Philosophical Remarks* and the *Philosophical Grammar*. The chief difference between the versions of the argument in these two works is the motivation for the claim that the meaning of a mathematical proposition is inseparable from its proof. In the *Philosophical Remarks*, Wittgenstein believes that the meaning of a mathematical proposition is inseparable from its proof because he believes that the meaning of a mathematical proposition *is* its proof. This claim, in turn, is backed by a general belief in verificationism: the meaning of any proposition is the means for showing its truth. In the *Philosophical Grammar* the meaning of a proposition is inseparable from its proof because the proof changes the meaning of the proposition. In

the *Philosophical Grammar* the meaning of a sentence is its role in the calculus in which it is used, and by proving a proposition one shows it to be a part of a different calculus than it was part of before. This distinction between the arguments of the two books is muddied by the fact that Wittgenstein had not fully abandoned verificationism by the time of the *Philosophical Grammar*. In each work the clarity of the argument is also damaged by the fact that the elements of the argument are often presented as freestanding assertions, so the relationship of support between them is not clear. We should look at each argument in detail.

The meaning of a sentence and the means of verifying it are linked throughout the *Philosophical Remarks*. Sometimes the link is weak. In these cases, the means of verifying a sentence merely act as a *constraint* on what the sentence can mean: “You cannot use language to go beyond the possibility of evidence” (§7), “It isn’t possible to believe something for which you cannot image some kind of verification” (§59). Sometimes, the means of verifying a sentence is a *part* of its meaning, “For, in a very important sense, every significant proposition must teach us through its sense how we are to convince ourselves whether it is true or false” (§148). Other times, the means of verifying a sentence are *equated* with the meaning of the sentence. “The meaning of a question is the method for answering it... Tell me *how* you are searching and I will tell you *what* you are searching for” (§27, cf. §§149 [6] and 150 [10]).

The strongest version of the verificationist thesis is necessary for the antiempiricist argument presented in *PR*. Sometimes the antiempiricist thesis is portrayed as following directly from a weak verificationism, “If you want to know what a proposition means,

you can always ask ‘how do I know that?’ Do I know that there are 6 permutations of 3 elements in the same way in which I know that there are 6 people in this room? No. Therefore the first proposition is of a different *kind* from the second.” (§114 [3], cf. §166 [3]). This clearly moves too quickly, however. After all, “There are six people in the room,” and “Protons are composed of quarks,” also have very different means of verification, but Wittgenstein would not want to say that they are of fundamentally different types. He must identify a significant difference in the means of verification. Wittgenstein does this by identifying the meaning of a mathematical proposition and its proof. “The completely analyzed mathematical proposition is its own proof” (§162 [2], cf. §§117 [2], 122 [5], 132, 148, 154 [4], and 161 [6]). This is a strong claim, and Wittgenstein can only pull it off if he has adopted the strongest version of verificationism. If the meaning of a statement is identical with its means of verification, then the meaning of a mathematical proposition is its proof. Once Wittgenstein has claimed that the meaning of a mathematical proposition is its proof, the needed contrast with empirical propositions comes naturally. Wittgenstein can now claim that as soon as one understands a mathematical proposition, one knows it to be true. This stands in marked contrast to experiments in empirical science, where one can know how to perform an experiment, without knowing how it will come out. If empirical experiments were like proofs, one would not know what one was doing until after it was done. This is what Wittgenstein is thinking of when he compares a mathematical expedition to a polar expedition: “How strange it would be if a geographical expedition were uncertain whether it had a goal, and so whether it had any route whatsoever.... But this is precisely

what it is like in a mathematical expedition. And so perhaps it's a good idea to drop the comparison altogether." (§161, cf. §§153 [1], 154 [8], and 190). The situation Wittgenstein presents to the mathematician is a lot like the paradox Meno presents to Socrates. How can we search for a mathematical theorem? If we know what we are looking for, we have already found it, and if we do not know what we are looking for, how can we search?

The idea that once one knows what a proposition means, one knows it is true is, I think, what is behind Wittgenstein's claims that "Arithmetic is its own application" (§109) and that "Arithmetic doesn't talk about numbers, it works with numbers" (ibid. cf. §157). If to know a mathematical statement is to know it is true, then one does not have to apply mathematical statements to something else to know that they are true. Moreover, we do not need to think of mathematics as describing an independent realm of numbers. We can understand mathematical truth as arising simply from *working with* numbers as tokens or inscriptions. However, although the argument I have reconstructed can explain many things Wittgenstein says, it is not necessarily the only way of ordering Wittgenstein's ideas in the *Philosophical Remarks*. Most of the themes I mentioned occur in many different contexts, and one could find many possible reasons that Wittgenstein could have advocated them. For instance, although he asserts many times that the meaning of a proposition is its proof, he only explicitly backs this idea by appeal to verificationism once. In §148 he writes, "'every proposition says what must be the case if it is true.' And with a mathematical proposition this 'what is the case' must refer to the way in which it is to be proved." I have taken this to be Wittgenstein's primary

motivation for believing that the meaning of a proposition is its proof because it makes the most sense. It should be noted, however, that contexts that link the thesis that the meaning of a proposition is its proof to verificationism are far outnumbered by contexts in which the thesis is just stated baldly.

The problem with the antiempiricist argument of the *Philosophical Remarks* that I have reconstructed is that it has a number of bizarre and undesirable corollaries, all of which Wittgenstein was aware of and seemed to embrace. The consequence of his argument that fascinated Wittgenstein most was the idea that there can be no unanswered questions in mathematics. In order to meaningfully ask a question, I must know how to go about answering it. But in mathematics, once I know how to answer a question, I have a proof, and thus the question is answered. From this Wittgenstein concludes that “only in our verbal language (which in this case leads to a misunderstanding of logical form) are there in mathematics ‘as yet unsolved problems’” (§159 [5]). The belief that there are no unanswered questions in mathematics is what lies behind slogans like, “the edifice of rules must be *complete*” (§154 [8]), and “there are no gaps in mathematics” (§158 [3]). What happens, then, when a new proposition is proved? A new proposition is not an addition to our system of mathematical knowledge, it marks the end of the old system and the inauguration of a new system, where all the signs have new meanings: “It’s impossible to discover rules of a new type that hold for a form with which we are familiar. If they are rules that are new to us, then it isn’t the old form” (§154 [8]). The idea that there are no unanswered questions in mathematics is only one of the strange consequences of Wittgenstein’s middle period antiempiricism. Another oddity is that

there cannot be two proofs of the same proposition (§153). If there are two distinct means of verifying a sentence, then the sentence has two distinct meanings. Also, the negations of true propositions are not false, they are senseless (§148). After all, there is no way to verify them. These strange consequences did give Wittgenstein pause. “My explanation mustn’t wipe out the existence of mathematical problems,” he worried (§148). But at this stage he showed no sign of abandoning his project because of them.

Writers in the mainstream of the philosophy of mathematics (e.g., Maddy 1986) tend to view the bizarre consequences of Wittgenstein’s antiempiricist argument in the *Philosophical Remarks* to be a reductio ad absurdum of his premise. It is easier to reject a strong version of verificationism than to accept the idea that there cannot be two proofs of the same proposition, or that mathematical questions are senseless, or that negations of proven mathematical statements are senseless. I agree with this analysis and would only add that Wittgenstein’s argument here goes against his own principles. Wittgenstein always portrays himself as removing confusions caused by philosophical misunderstandings of language. The *Philosophical Remarks* opens with just such an avowal. He claims to be interested in clarifying the grammar of ordinary language (§§1, 3), and that previous philosophers have perpetuated confusions because they have failed to understand the grammar of ordinary language (§9). But in his antiempiricist argument, he draws a conclusion that flies in the face of common sense, on the basis of a philosophical theory of meaning. Rather than ending philosophical theorizing by clarifying grammar, he has given us an example of philosophy of the most theoretical sort.

By the time of the *Philosophical Grammar*, Wittgenstein had developed a more complex relationship to verificationism. Part I of the published version of the *Philosophical Grammar* opens with an outline of a fundamentally new approach to meaning. The meaning of a word is constituted by the rules that govern its use. Such rules taken together form a grammar. There are grammars for many different uses of language—a grammar for describing sensations, a grammar for mathematics, etc. The meaning of any particular word is therefore a function of the grammar it belongs to: “The place of a word in grammar is its meaning” (I §23 cf. I §§14, 27, 31, 84). Sometimes this new conception of meaning is stated more broadly, as in “the use of a word in language is its meaning” (I §23, cf. I §29). These passages emphasize the idea that a word gets its meaning from being used a certain way, from playing a role in a person’s life. Passages like this fit better with the conception of language of the *Philosophical Investigations*, where language is seen as something more smoothly integrated into people’s lives as a whole.

Although the first half of the published *Philosophical Grammar* offers a new understanding of meaning, the second half repeats essentially the same argument against mathematical empiricism that was proffered by the *Philosophical Remarks*. Indeed, a superficial reading of the second half indicates that he is relying on the same verificationist understanding of meaning used in the *Remarks*. This could be explained as an artifact of the way the *Philosophical Grammar* was edited. The published *Philosophical Grammar* is an amalgamation of three sources, written over a period of time when Wittgenstein’s philosophical views appear to have been changing rapidly. The

primary source is the so-called “Big Typescript,” a collection of remarks, entitled the *Philosophical Grammar*, which he dictated to a typist with the intent to publish it. Part II of the published *Philosophical Grammar*, which contains his antiempiricist arguments, comprises the second half of the Big Typescript. Part I of the published *Philosophical Grammar*, however, was assembled out of revisions Wittgenstein made to the Big Typescript in the year or so after he dictated it. These revisions come from two sources, a manuscript marked revision from 1933, and loose folio sheets that could have been written as late as 1934. Wittgenstein’s new ideas on meaning come from these revisions, primarily the later loose folio sheets. It might be reasonable to think that the second half of the book is more verificationist than the first because it was written earlier, before his break with verificationism.

This would be a mistake, however. By the time he wrote the Big Typescript, Wittgenstein had already been discussing the ideas about meaning that are found in its later revisions. The *satzsystem* conception of meaning, as his new view is called, is present in Wittgenstein’s conversations with Schlick and Waismann, as recorded in *Wittgenstein and the Vienna Circle*. (See, e.g., *WWK*, p.66.) Other conversations and lectures indicate that that Wittgenstein felt that the *satzsystem* conception of meaning and some sort of verificationism were completely compatible. In some of his Cambridge lectures, for instance, he endorses both the *satzsystem* view and verificationism, as if they were aspects of the same thing: “If you want to know the meaning of a sentence, ask for its verification. I stress the point that the meaning of a symbol is its place in the calculus, the way it is used.” (*AWL*, p.29). The reason verificationist language appears alongside

statements of a *satzsystem* conception of meaning is that after the adoption of the *satzsystem* view of meaning, Wittgenstein continued to hold a weakened, benign verificationism that was compatible with his new approach.¹¹ According to his weakened verificationism, identifying the method of verification of a sentence was a good way to determine what *satzsystem* the sentence belonged to. The importance of verification, however, was purely epistemic, and this is reflected in the language Wittgenstein uses in the second half of the *Philosophical Grammar*, where everything is put in terms of how to discover the meaning of a proposition. For instance, §24 of Part II, the section that is crucial to the antiempiricist argument of the *Grammar*, is entitled “If you want to know *what* is proved, look at the proof.” When Wittgenstein wants to discuss what meaning actually consists of, he switches to language more compatible with the *satzsystem* conception of meaning. For instance, in §24 Wittgenstein writes, “behind the words ‘I know ...’ there isn’t a certain state of mind to be the sense of those words. What can you do with that knowledge? That’s what will show what the knowledge consists in” (*PG* II §24 [12]). Shortly thereafter, he writes “whether a pupil *knows a rule* for ensuring a solution to $\int \sin^2 x \, dx$ is of no interest; what does interest us is whether the *calculus* we have before us (and he happens to be using) contains such a rule” (*PG* II §25 [10]).

Now, this weakened form of verificationism clearly isn’t sufficient to support the antiempiricist argument of the *Philosophical Remarks*. That argument depended on the claim that it was impossible to characterize a mathematical proposition without actually

¹¹ Shanker (1987, ch. 2) presents a similar account of the relationship between verificationism and the *satzsystem* conception of meaning.

proving it. However, if checking the verification of a proposition is just a method for determining what *satzsystem* a proposition belongs to, it would be perfectly possible to characterize a proposition without proving it. One merely has to find other means to place it in the context of a given *satzsystem*. The version of the middle period argument that appears in the *Philosophical Grammar* takes another route to the contrast between empirical and mathematical propositions. The key idea for this version of the argument is that when a proposition is proved, it is placed in the context of a mathematical calculus that it was not in before. Once Gödel proved completeness for the first-order predicate calculus, the theorem had a place in the first-order predicate calculus that it did not have before. Thus Wittgenstein writes, “A mathematical proof incorporates the mathematical proposition into a new calculus, and alters its position in mathematics. The proposition with its proof doesn’t belong to the same category as the proposition without its proof” (*PG* I §24, [9]). One might object that completeness always belonged to the first-order predicate calculus, and we just didn’t know it. But, Wittgenstein would reply, if we didn’t know it, our use of the word “completeness” was not governed by the first-order predicate calculus. Indeed, all the meaning the term had was the vague meaning ordinary language could give it. This is why Wittgenstein refers to unproved propositions as “signposts for mathematical investigation, stimuli to mathematical constructions” (*ibid.*). Thus again mathematical statements are unlike empirical statements because the proposition one proves cannot be stated until it is proved. “To put it concisely,” Wittgenstein writes, “the mathematical proof couldn’t be described before it is discovered” (*ibid.*).

This new version of the middle period antiempiricist argument has all of the unreasonable consequences that the first does. Indeed, some of the most unreasonable consequences—that one never proves the proposition one sets out to prove and that unproved propositions are strictly meaningless—appear as necessary steps of the argument. Wittgenstein's willingness to consider such ideas shows more than just his ability to challenge accepted dogma. It shows his belief that extreme measures were necessary to rid ourselves of philosophical confusion. Nevertheless, I feel comfortable saying that if no one can prove what they set out to prove, something must be wrong here. As with the previous version of this middle period antiempiricist argument, I think the problem is that Wittgenstein has gone against his own best instincts. His argument depends on the idea that a proposition gets its meaning from being part of a calculus that exists in radical isolation from other calculi, so that the role a proposition plays anywhere else cannot affect its meaning in a given calculus. His bizarre conclusions stem from sticking to this philosophical theory even when it challenges common sense. The impasse he is in will be solved when he moves to an even more holistic conception of meaning. Part of the transition from the middle to the late Wittgenstein involves the realization that a mathematical calculus is smoothly integrated into the rest of a person's life. This realization is marked by his abandonment of all examples from advanced mathematics and his focus on extremely simple activities. But once we realize that a calculus does not exist in radical isolation, there is no need to view a proposition that has been proved to be a part of the calculus as of a different kind than an unproved proposition. Thus the move

to the later stage of his life marks a move away from the flawed arguments of the *Philosophical Remarks* and the *Philosophical Grammar*.

The Late Period Argument

The late period in Wittgenstein's work is marked by his attention turning to examples of individuals attempting to follow simple rules, such as the student who attempts to follow the rule 'add two'. The fact that Wittgenstein's last argument against mathematical empiricism was written as an extension of his thoughts on rule following shows that this argument marks a substantial break with his previous antiempiricist arguments. Wittgenstein's late period argument against mathematical empiricism is presented in the collection of fragments that now forms the first book of the *Remarks on the Foundations of Mathematics*. This collection is a well-organized whole, written between 1937 and 1938 and worked over extensively up until 1944. It was originally meant to form the second part of the *Philosophical Investigations*, to be placed right after remark 188 and Wittgenstein's famous discussion of the pupil who has trouble counting by twos (see the editor's introduction to the *RFM* and Monk 1990, 380). The work begins with a continuation of the discussion of rule following in the *Philosophical Investigations* §§139–88, moving on to apply these insights to the philosophy of mathematics. As I see it, the initial remarks on rule following are meant to show that the normativity that governs all language is irreducible to any other feature of language or nature. I understand this part of the argument as beginning in the *Philosophical Investigations* §§138–88 and continuing through *Remarks on the Foundations of Mathematics* §§1–23.

The remaining parts of the *RFM* extend these thoughts into an argument for Wittgenstein's two antiempiricist theses—that mathematical statements are not descriptions and that they are expressions of norms. Wittgenstein's argument for the first thesis works by showing that mathematical operations have a property I term *result dependence*: The result of a mathematical operation is a part of the definition of that operation. So, for instance, it is a part of the definition of collecting two pairs of objects that the resulting collection has four objects. Any other result, and you have not really collected two sets of two. As we shall see, result dependence means that there is simply no room for empirical results to affect mathematical reasoning. Wittgenstein's argument for the first antiempiricist thesis is an outgrowth of his thoughts on rule following: he first shows that norms are irreducible, then he shows that mathematical results are determined entirely by the rules governing mathematical operations. Once result dependence and the first antiempiricist thesis have been established, the second antiempiricist thesis can be brought in as an alternative account of the nature of mathematical statements. Let's look at the parts of Wittgenstein's late period argument in order.

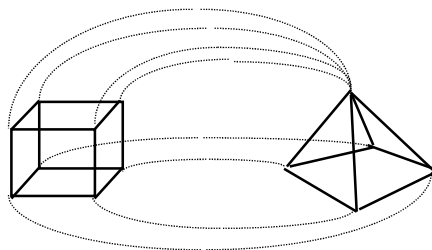
Wittgenstein's inquiry into rule following in the *Philosophical Investigations* begins by examining the "a-ha" experience, where one suddenly understands a proposition or an

idea that had previously been mysterious.¹² In such a circumstance we feel as though we immediately grasp the rule that governs the use of the concept or gives us the meaning of a term. All at once, we see what uses of the term are correct and what uses are incorrect. The thrust of Wittgenstein's remarks on rule following is that focusing solely on this kind of experience can give us a very misleading picture of rules, and hence of meaning. If we take the a-ha experience as a paradigm for understanding a rule, we will be tempted to view the rule as a kind of super-entity, capable of judging the correctness of every future application of itself and yet being stuffed into our heads all at once. Wittgenstein's treatment of rule following shows that nothing can stand above the applications of a rule and determine which applications are correct and incorrect, least of all something that could fit into our heads. The conclusion we are left with is that a rule simply *is* its correct

¹² I should explain how my account of Wittgenstein's rule-following argument relates to other commentator's expositions, not because I need to position myself in the ongoing debates among Wittgensteinians, but because I need to acknowledge my intellectual debt. In general, I owe my understanding of Wittgenstein's ideas on rule following to the writings of Meredith Williams (1983/1999, 1991/1999, 1994a/1999, 1994b/1999, and Forthcoming/1999, in particular) and his ideas on the philosophy of mathematics to Stuart Shanker (1987). By stating Wittgenstein's rule-following arguments in the form of a reductio, I am favoring the interpretation of Shanker (1987, 13 ff.) over Kripke (1982). By emphasizing that it only *seems* as though we are left without a way to extend a pattern according to a rule, I am challenging the antirealist interpretations of Wittgenstein given by Dummett (1959/1978) and Wright (1980). On the other hand, by saying that a rule simply is its correct application, I believe I am disagreeing with Baker and Hacker (1980 and 1985, 1984). Their statement that a rule bears an internal relationship to its applications strikes me as a weaker claim. Finally, I should note that for now I am remaining silent on the issue of whether Wittgenstein's thoughts on rule following imply that one can only follow a rule in the context of a community of rule followers. It is not necessary to take sides on this issue in order to elaborate Wittgenstein's critique of empiricism. I will turn to it, however, in the last chapter, when I outline my positive views.

applications. There is no rule apart from the right and wrong actions. But this does not mean that there is no right or wrong in language, far from it. In fact, it is the other understandings of a rule that take the normativity out of language, because they reduce the normativity of language to something else by trying to explain right and wrong acts in terms of something other than right and wrong acts. Wittgenstein leaves normativity irreducible.

The first vision of a rule as super-entity that Wittgenstein addresses is the image of a rule as a picture that we compare to actions in the world to determine if those actions are correct or incorrect. Suppose, he suggests in *PI* §139, that when I grasp the rule for applying the word ‘cube’, an image of a cube appears in my mind’s eye. I can then say that I apply the word ‘cube’ correctly if I apply it to objects that look like this cube. But is this image sufficient to divide the uses of the word ‘cube’ into correct and incorrect? Wittgenstein, in one of the most famous moves in the *Philosophical Investigations*, illustrates that it is not. What is to prevent someone from applying their image of a cube like this:



and saying that a cube is a pyramid? To produce a division of right and wrong applications one needs both a mental cube and a *scheme of projection*. But if a scheme of projection is needed then there is no limit to what we will need to apply our image of a cube. How are we going to understand our scheme of projection? “Can’t I now

understand different applications of this scheme too?” (*PI* §142). Another scheme will be required, and that scheme will need a third, etc. Thus Wittgenstein illustrates by means of a *reductio ad absurdum* how a mental image is insufficient to fix the nature of a rule. If we assume mental images allow us to apply a rule, an intolerable regress results. The regress argument also has the same form (although possibly not the same intent) as the argument in Lewis Carroll’s popular essay “What the Tortoise Said to Achilles” (1895).

Following Meredith Williams (1991/1999) and Robert Brandom (1994), I will distinguish the regress argument found in *PI* §§139–41 from a companion argument, which Williams dubs the *paradox of interpretation*, and Brandom the *gerrymandering argument*. In this argument we are again asked to consider radical possibilities for understanding a rule. This time the occasion is the instruction of a pupil in how to write a sequence of numbers. The situation is first mentioned in the passages immediately following the regress argument, §§143–45. Wittgenstein begins by considering the ways we might explain to someone how to write the natural number sequence, for instance by having her copy the numbers 0–9, guiding her hand as she writes. Wittgenstein immediately cuts short this scenario, pointing out that “here already there is a normal and abnormal learning reaction.” (*PI* §143) The instructions one gives here can be misinterpreted, just as the cube was misprojected. The pupil, when asked to write the numbers 0–9 on her own, may write any series at all and claim it was a valid extension of the examples she was led through. The point of mentioning this possibility here is the same as it was in the regress argument: to separate the action of rule following from any entity that might stand behind it. Here, however, the emphasis is placed on the

multiplicity of relationships between the entity that is supposed to guide our actions and the actions themselves, rather than on the multiplicity of guides one might call in to try to save rule following. The emphasis on the multiplicity of relations between rules and actions is what Brandom is thinking of when he labels this the gerrymandering argument. The argument here also differs from the regress of interpretation argument in its generality. While the regress argument only dealt with interpretations, this argument applies to *any* entity that might be used in training and might be thought to guide later action. This generality is what elevates the argument to the status of a paradox, although its paradoxical nature is not brought out until some time after the argument is first introduced. After the discussion of interpretation in §§143–45, Wittgenstein pauses to explore other themes—counterpoints and harmonies to §§143–45. In §185 he returns to the example in §143, this time considering a student being taught to count by twos. The pupil again understands the teacher incorrectly. He counts rightly up to 1000, but afterwards counts 1004, 1008, 1012, Now Wittgenstein imagines the teacher reaching complete exasperation. Nothing he can say will change the student’s understanding. “It would now be no use to say ‘but don’t you see!’—and repeat the old examples and explanations—in such a case we might say perhaps: it comes natural to this person to understand our order with our explanation as we should understand the order ‘add 3 up to 1000, 4 up to 2000, 6 up to 3000 and so on’” (*PI* §185). The unacceptable consequences of the gerrymandering argument now appear to be completely general: there is no way that a rule can determine its applications. Wittgenstein expresses the seeming lawlessness of extending the series of even numbers in the next remark, when he writes, “It would

almost be correct to say, not that a new intuition was needed at every stage, but that a new decision was needed at every stage” (*PI* §186). But Wittgenstein backs away from this conclusion, pointing out that there is still an assumption that has led us to this state. When we first began this venture in *PI* §139 we were looking for a way to view a rule as a super-entity capable of determining its applications in advance. The idea of a rule for the word “cube” as a mental image was one attempt to understand a rule as a super-entity. Our current attempt to understand how people are trained to follow rules still tries to view rules as super-entities—we want the series of even numbers after 1000 to be contained within the series of even numbers before 1000. Wittgenstein initially wrapped up this section of the argument by reminding the reader of the crucial premise that brought us to this position: “Here I should first of all like to say: your idea was that the act of meaning the order had in its own way already traversed all those steps: that when you meant it your mind as it were flew ahead and took all the steps before you physically arrived at this or that one” (*PI* §188). Later, after he had abandoned Book 1 of the *RFM* as a part of the *PI*, he wrote this famous summary of the gerrymandering argument: “This was our paradox: no course of action could be determined by a rule, because every course of action can be made out to accord with the rule. The answer was: if everything can be made out to accord with the rule, then it can also be made out to conflict with it. And so there would be neither accord nor conflict here.” (§201). As difficult as it may seem, we must reject the idea that a rule exists beyond its correct application. Super-entities can neither accord nor conflict with actions. Thus nothing can stand behind the correct application of a rule, determining its correctness. There is nothing but the application.

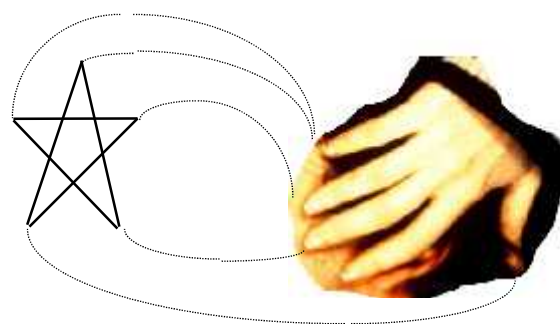
This does not mean, however, that these actions cannot be right or wrong. To deny that there are right and wrong actions would be to fall into linguistic nihilism. But if we are to recognize some actions as according with a rule and others not, and yet deny the existence of super-entities, we must regard correctness as an irreducible property of the action. This is why Wittgenstein claims that “there is a way of grasping a rule that is *not* an *interpretation*, but that is exhibited in what we call ‘obeying a rule’ and ‘going against it’ in actual cases” (*PI* §201). There is nothing beyond the right and wrong action that one grasps when one follows a rule. The normativity of rules is irreducible.

The passages that begin Book 1 of the *Remarks on the Foundations of Mathematics* concur with the evaluation of the gerrymandering argument given in the final version of the *Philosophical Investigations*. Sections 1 and 2 are basically identical to *PI* §§189 and 190, the chief difference being that the mathematical examples used in the *Remarks on the Foundations of Mathematics* have been simplified in the more polished *Philosophical Investigations*, another instance of the later Wittgenstein’s desire to use only commonplace examples. The subsequent passages clarify different aspects of the work on rule following he has done up to that point. He emphasizes that he does not believe that anyone would actually misinterpret the instruction “add two” (*RFM* I §3) and deals with the objection that he cannot explain the inexorability of mathematics (*ibid.* §§4–5). Sections 6–23 argue that inference cannot be explained as a mental event.

Starting in *RFM* I §24, Wittgenstein applies these thoughts on rule following to the thesis that concerns us here—mathematical empiricism. He begins this section by setting a challenge for himself: “Separate the feelings (gestures) of agreement, from what you *do*

with a proof.” (*RFM* I §24). His thoughts on rule following have shown that the feelings and images that accompany words have nothing to do with their meaning—this was the upshot of the regress of interpretation. On the other hand, Wittgenstein is now firmly established in his belief that what one actually does with a piece of language is its meaning. Therefore, by saying that he wishes to separate the feelings that accompany a proof from what one actually does with it, he is announcing that he wishes to identify the real meaning of a proof, the meaning that it has when philosophical illusions have been swept away. The discussion that immediately follows this announcement (§§25–35) neatly encapsulates what he wishes to say about mathematics. This section focuses on a familiar sort of problem: how to tell when two figures represent the same number, e.g., a hand and a pentacle. The obvious way to determine equinumerosity here is to draw lines between the fingers of the hand and the vertices of the pentacle. But this operation can be conceived of in two ways. If we think of it as an operation done on just these two figures, then it shows nothing mathematical. But we can also view the operation generally, as a proof about certain kinds of figures, which Wittgenstein labels H (for hand) and P (for pentacle). At this point the operation becomes mathematical, and, as Wittgenstein points out, atemporal. Many different acts of putting objects in one-to-one correspondence are compressed into a single figure. The drawing of lines running between a hand and a pentacle, “now serves as a new prescription for ascertaining numerical equality: if one set of objects has been arranged in the form of a hand and another as the angles of a pentacle, we say the two sets are equal in number” (*RFM* I §30). The word “prescription” [Vorschrift] is chosen carefully. The drawing does not just illustrate a workable method

for determining equality; it tells us that we must always consider collections that can be arranged this way to be equal. In the next passage Wittgenstein's anonymous interlocutor objects that it would not be possible to draw lines between collections of types H and P in such a way that the sets would come out unequal. Therefore there is no need for a prescription here. There is nothing being ruled out. Wittgenstein replies that there are drawings that one might take as showing that H and P are unequal, e.g.,



Of course, if we were to produce such a drawing we would assume that we had connected the hand and the pentacle wrong. But this is just because there is a prescription at work here. The fact that we rule out such alternative ways of producing correspondences also shows that when we first produce the correspondence between the hand and the pentacle we are not really comparing the two figures. Really what we are doing is picking out one of the many possible pairs of hands and pentacles and saying that it is representative of the kind of pair that gets called “correct.” In this vein, Wittgenstein writes, “The proof doesn’t *explore* the essence of the two figures, but it does express what I am going to count as belonging to the essence of the figures from now on.—I deposit what belongs to the essence among the paradigms of language,” and then on the next line he concludes, “The mathematician creates *essences*.”

The thought experiment of the hand and the pentacle is a compressed statement of everything he has to say in the remainder of the *Remarks on the Foundations of Mathematics*. As I see it, five themes are at work in this passage.

1. Proofs must use exemplars of objects or types of objects, rather than particular objects.
2. We reject actions and experiences that do not conform to our mathematical ideas.
3. Mathematical operations are result dependent.
4. Mathematical statements are not descriptions of the world.
5. Mathematical statements are expressions of norms.

The first two themes are basically observations about ordinary mathematical practice. The first came up when Wittgenstein noted that to make the picture of the hand and pentacle into a mathematical proof, we had to view the individual hand and the individual pentacle as representations of certain classes of figures. The second theme came up when Wittgenstein pointed out that we reject ways of connecting the hand and pentacle that show that they are not equinumerous. The remaining three themes, I think, are conclusions drawn on the basis of the first two themes. The third theme is the idea that mathematical operations are result dependent, that is, their results are included in their definition. In this case, the definition of the operation of putting the hand and pentacle in one-to-one correspondence includes getting the result that there are the same number of points in each. The next theme is one of Wittgenstein's two main antiempiricist theses. It comes up in this passage when Wittgenstein says that "The proof doesn't *explore* the essence of the two figures." As we shall see, this thesis is actually a logical consequence of result dependence. The final theme is Wittgenstein's other antiempiricist thesis, the claim that mathematical statements are expressions of norms. It comes up in this passage

with Wittgenstein's extensive talk of proofs as prescriptions. This theme is also sometimes expressed with the idea that a proof is a "memorable picture." The memorable picture idea comes up here in the form of Wittgenstein's insistence that the hand and pentacle form a single figure, and that an important feature of this figure is that it is simple and easy to grasp. It must be simple and easy to grasp because it has to serve as an expression of a norm. Exactly how it is supposed to do this is unclear, though. In *RFM* I §§25–35 Wittgenstein seems to think that the hand and pentacle together serve as a model used to judge future correlations. In this frame of mind, Wittgenstein seems rather at odds with himself. The proof appears to be a model for future action in almost the same way that the picture of a cube served as a model for the meaning of the word "cube" in the *Philosophical Investigations*. I would like to propose a different way of understanding the expressive role of mathematical statements. Instead of serving as a model for future correlations, we can understand the hand and pentacle as expressing the nature of the class of correlations that would be called correct whether or not we chose to explicitly state the norm uniting them. Earlier in this chapter, I distinguished two ways a statement could be an expression of a norm. It could either be the founding of a new norm, like a New Year's resolution, or it could express a rule that one was already implicitly committed to following. When Wittgenstein treats proofs as memorable pictures, he is effectively saying that mathematical statements are expressive of norms in the way that New Year's resolutions are. I, on the other hand, would like to treat them as expressive of pre-existing norms. This way of understanding the expressive role of mathematics is parallel to Brandom's expressivist view of logic.

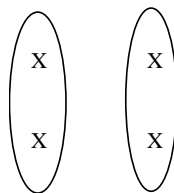
These five themes are explicated more fully in the subsequent parts of Book 1. There they are joined by three other important themes—the role of imagination in proof, the impossibility of universal error in mathematics, and the image of a proof as unfolding an essence. These new themes, like the first two themes in the hand-and-pentacle passage, serve as premises, while the last three themes in the hand-and-pentacle passage continue to serve as conclusions. The list of five themes on p. 92 would thus be renumbered like this:

1. Proofs must use exemplars of objects or types of objects, rather than particular objects.
2. We reject actions and experiences that do not conform to our mathematical ideas.
3. Imagination plays a role in mathematics that it does not play in the sciences.
4. Widespread universal error is impossible in mathematics.
5. Proofs are somehow seen as unfolding an essence.
6. Mathematical operations are result dependent.
7. Mathematical statements are not descriptions of the world.
8. Mathematical statements are expressions of norms.

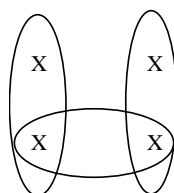
To bring out the ways in which I think the argument of the book is persuasive, I'm going to present these ideas in a rather un-Wittgensteinian manner. I am going to portray themes 1–5 as pieces of *data* about mathematical practice. I will present theme 5 as an *explanation* of these data. Theme 7, the claim that mathematical statements are not descriptions, will then follow from theme 6, the claim that mathematical operations are result dependent. Once Wittgenstein establishes that mathematical statements cannot be descriptions, some other explanation of their function is called for. The need for this explanation motivates our acceptance of theme 8.

This presentation does some violence to Wittgenstein’s conception of his own work. The individual themes can be seen as separate descriptions of different contrasts between mathematical practice and experimental practice. If Wittgenstein really thought he was offering nothing but descriptions, this may be the way he viewed his own work. On such a reading, earlier themes would only serve to introduce later themes—perhaps to soften the reader up for the more radical claims. However, the themes I am labeling “data” are not very convincing on their own, and the other two themes serve remarkably well as explanations. The work is simply much stronger as a whole if we posit a stronger, explanatory relationship between the earlier themes and the later one.

The idea that mathematical proofs cannot be about particular figures came up first when Wittgenstein replaced talk of the individual hand and pentacle with the symbols ‘H’ and ‘P’. This theme is raised again in §38. In that passage Wittgenstein compares two ways of going from a particular figure to a universal mathematical statement. He points out that if one can use four Xs in a figure like this



to show that two plus two equals four, then one can easily use the same four Xs like this



to show that two plus two plus two equals four. The example serves to reinforce the idea that mathematical propositions simply cannot be about particular objects. There is no fixed way to go from a specific figure to a general proposition. If the mathematical proof does not start as a representative of a kind, it cannot get there later. Another look at this same idea comes in §§55–57. There his example is a geometric proof involving the construction of a rectangle. In §55 he asks why constructing one rectangle was sufficient for the proof. If the construction were an experiment, surely the sample is too small. In §57, he then asks us to consider a more restricted proposition, namely, that the proof shows something about the specific rectangle in the drawing. If the proof were an experiment, this more restricted proposition would be amply demonstrated. But the restricted proposition is not a proposition of geometry at all. Remarks §§55–57 can be seen as showing that the effective mathematical proof is not a successful experiment because the sample is just too small, while §38 shows that the successful experiment cannot be a mathematical proof because there is no fixed way to go from the particular to the general.

The theme that resonates most closely with Wittgenstein's rule-following arguments is the image of possible events being rejected on the basis of mathematical ideas. We saw this in the hand-and-pentacle example when the alternate ways of linking the hand and pentacle were rejected. Another example occurs in *RFM* I §37, where Wittgenstein asks us to imagine placing a pair of apples on a table and then placing a second pair next to them. If arithmetic statements were empirical, we could imagine this as a sort of experiment designed to determine if two plus two really did equal four. But, Wittgenstein

points out, suppose when we tried the same experiment with beans we found that we sometimes only got three beans. We would not conclude that two and two were not always four. We would say that one of the beans disappeared. The proposition “two plus two is four” is used here to judge whether we have collected beans correctly. In §87 Wittgenstein makes the same point with regard to an attempt to demonstrate that $10 \times 10 = 100$ by arranging 100 marbles in 10 rows of 10. (Wittgenstein archly refers to this as “doing the drill with marbles.”) Wittgenstein notes that if we kept getting different results when we did the drill, we would say that, at least some of the time, we were arranging the marbles incorrectly. In all of these examples an event in the world that might falsify a mathematical statement—an attempt to match the fingers of a hand and the points of a pentacle that shows they do not match, an attempt to collect two sets of two objects that yields three objects total, an attempt to create a ten-by-ten array of marbles that does not yield 100 marbles—is declared to be impossible. And in each case the reason for this impossibility is that some action is judged to have been performed incorrectly.

On their own, examples of the types found in these two themes do not prove much. They are, in fact, perfectly compatible with mathematics as an empirical science. This is clearest in the case where mathematical ideas are used to judge the correctness of an experiment. If “two plus two equals four” were merely an incredibly well-confirmed empirical proposition, it could still be used as a check on observations, such as the case where we put two pairs of beans on a table. There are plenty of cases where an experiment is assumed to have gone wrong if its results conflict with better established

results. The examples of generality in mathematics similarly have parallels in empirical science. There are cases in empirical science where an event need only occur once to prove a theory. For instance, Peter Galison in *How Experiments End* (1987) describes the role of such “golden events” in convincing experimentalists of the existence of muons and, later, weak neutral current. Moreover, it is not even always the case that in mathematics we always consider a proposition true after one proof has been produced. If some of the steps in the proof are controversial, or if the proof is too long to be surveyed easily, mathematicians will sometimes withhold judgment until a second proof has been produced using other means.¹³ To the extent that there is a difference in the number of verifications mathematics and the sciences require of a proposition, the empiricist can explain it. If we assume that the methods of mathematics are methods of experimentation whose accuracy is extremely well verified, it would follow naturally that most of the time a mathematical experiment need only be performed once. Another part of Wittgenstein’s treatment of the theme of generality was the claim that an individual mathematical figure, such as his collection of four Xs, can be interpreted in too many different ways to be informative. But this is simply an instance of a well-known aspect of experimentation: the underdetermination of theory by evidence.

Fortunately, there are more aspects of mathematical practice Wittgenstein wants us to consider. One significant feature of mathematical proof is that physical objects need to be involved in them at all. Wittgenstein first brings this out in §36 in reference to the drill with marbles. Wittgenstein points out that one could make a film of someone doing the

¹³ I owe these last two points to Arthur Fine.

same drill and the proof would be no weaker. The converse idea, that one cannot perform true experiments in the imagination, is discussed in §§96–98, where he compares the process of examining a tangent to a curve drawn on a piece of paper to the process of imagining the same tangent. One can examine smaller and smaller sections of a visual curve to find a point where the segment appears straight and a tangent would intersect the curve everywhere in the segment. One cannot do something similar for an imagined curve, and not simply because a line of constant curvature is defined as intersecting a straight line at a point. One cannot perform an analogous experiment with an imagined curve because there are no criteria for judging when an imagined curve begins to seem straight. From this Wittgenstein concludes “I can calculate in the medium of imagination, but not experiment.” (*RFM I* §97).

At first, this datum may not strike you as being true at all, or if it is true, it is only true for simple calculations. Most people can calculate with small numbers in their head, perhaps doing a drill with three rows of three marbles. But ten rows of ten or a hundred rows of a hundred is certainly beyond our ken. This is actually an objection that can be raised about many of Wittgenstein’s mathematical observations. Wittgenstein’s insistence on using only the simplest examples seems to skew his perspective. My general response to this kind of objection is to assert that there is no great ontological divide along the number line. If we conclude on the basis of simple operations that the imagination plays a role in calculation that it does not play in experiment, then this must also hold for more complicated operations. There is, at least in principle, a way we can use the imagination there that is not open to us in experiment.

Another aspect of mathematical practice Wittgenstein points out is that it is inconceivable for everyone to be consistently wrong in a calculation (*RFM* I §§135–37). If we consistently got some other answer to a mathematical problem, we would simply declare that to be the new correct answer. The calculation with a different, allegedly “correct” answer would simply be called another function with the same domain. For instance, if we consistently found that 12×12 was 145, we would adopt this as our new definition of “ \times ” and use a new symbol for the notion of multiplication that yields the other “correct” results. In a way, the impossibility of universal error in mathematics plays on the same flexibility in the definition of mathematical functions that the example of adding two plays on. The example of adding two depended on the fact that given any finite string of numbers, one can always define a function that will make any additional number the correct next step in the sequence. Similarly, the impossibility of universal error depends on the idea that there is a function that corresponds to any pattern of mapping pairs from ordered pairs to real numbers, so long as that mapping is consistently applied, so that 12×12 is not sometimes 144 and sometimes 145.

This again may seem like an observation that only applies to small numbers and simple calculations. While it may be the case that we could not consistently calculate 12×12 incorrectly, surely there are errors in more elaborate calculations that can and have gone unnoticed for long periods of time.¹⁴ Before we put too much stock in this objection, though, we need to think about exactly what a counterexample to Wittgenstein’s observations would look like. It would not be enough for a mistake to go

unnoticed simply because no one checked it. There are many such cases, to be sure, but they are not instances of the same *calculation* being performed repeatedly, but incorrectly, by large numbers of people. It would not even be enough for an erroneous calculation to be relied on for a long period of time, because again an error is not being made repeatedly. What is needed is for the mathematical community to consistently calculate a function one way for many hundreds of years, and then switch to calculating it another way. In fact, there are plenty of examples of precisely this phenomenon. There are many things that are true in Leibniz's version of the calculus that are no longer true in the refined calculus produced by Cauchy and others in the nineteenth century. But this is not a case of a function being calculated incorrectly for many numbers of years. One function has been replaced by another function. Even if you regard the Cauchy definition as objectively better than the Leibnizean definition, and even if you believe that there is one subject matter, the nature of infinitesimal calculation, which both functions are attempting to capture, it remains the case that there are two functions being calculated here. And this is the phenomenon that Wittgenstein is pointing to when he says that widespread consistent error is impossible in mathematics.

There is a phenomenon related to the impossibility of widespread consistent error in mathematics that Wittgenstein did not note, probably because it requires some knowledge of the history of mathematics to understand, and Wittgenstein wanted to avoid all entanglements with other branches of study. I'm thinking of the phenomenon Michael Resnick (1997) has dubbed the "Euclidean rescue." When non-Euclidean geometries

¹⁴ Ed Avrill stressed this objection to me.

were developed, Euclidean geometry did not become false. Instead, it was given a new domain, Euclidean space. There are many cases in the history of mathematics of theories being “rescued” in this fashion. Wittgenstein’s claim that widespread consistent error is impossible in mathematics is equivalent to the claim that a Euclidean rescue is always possible in mathematics in a way that it is not possible for empirical science. One might think that is not really true. Can’t we perform a Euclidean rescue for Newtonian mechanics, saying that it is actually true of a new entity, Newtonian space? There is one key difference between this case and the Euclidean space. A rescued physical theory like Newtonian mechanics goes from being about a physical object to being about a mathematical object, if there is an object that it is about at all. A rescued mathematical theory does not change domains like this. It goes from being about one kind of mathematical object to another. This is a difference we should account for.

The final recurring image in the *Remarks on the Foundation of Mathematics* is the idea that mathematical proofs and calculations “unfold” properties of numbers that were already inside them. Wittgenstein describes the sense that calculation unfolds properties of numbers as the sense that one is “demonstrating an *internal property* (a property of the *essence*)” (§99). Now, it would be question begging to assume that there actually is such an essence that one unfolds in proof. However, it is true that many people have the sense that there is such an essence, and that is the datum Wittgenstein wants to explain.

These sorts of comments on their own are still little more than a collection of oddities about a particular science. However, there is a far more radical lesson Wittgenstein wants to draw out of these comments. Wittgenstein’s conclusion to remark §86—where he

pointed out that if we did not always find we could arrange 100 marbles into 10 rows of 10, we would assume that something had gone wrong during the process of arranging—was that, “this shows that you are incorporating the result of the transformation into the kind of way the transforming is done.” In this passage, Wittgenstein moves from the idea that mathematics can be used to judge experience to the idea that mathematical operations are result dependent. Unfortunately, it is not clear how he thinks this follows. It doesn’t follow purely logically, since, as we have just seen, the role of mathematics in judging experience is compatible with the hypothesis that math is empirical. If we wished to be uncharitable, we might say that Wittgenstein’s arrangement of ideas is purely rhetorical. The remarks about the drill with marbles were just meant to soften the readers up before they encounter the radical theses. But I do not want to take this stance toward Wittgenstein’s work. I think the best alternative is to say that result dependence is an *explanation* of the judgmental role of mathematics. One possible reason that we say we must have arranged the marbles incorrectly when we fail to find 100 is that our definition of arranging marbles is result dependent. The competing explanation from the empiricist would be that $10 \times 10 = 100$ is an extremely well-confirmed explanatory hypothesis. We can compare Wittgenstein’s explanation and the empiricist explanation by examining their explanatory power. From what we have just seen, we know that both hypotheses are able to explain the fact that if a 10×10 arrangement of marbles did not contain 100 marbles, we would say something had gone wrong in the arranging. What of the other aspects of mathematical practice Wittgenstein has taken as his themes? On the theme of generality, the two hypotheses are also on a par. We have already seen how the empiricist

hypothesis can explain any difference between the number of proofs mathematics requires of a proposition and the number of experiments required to confirm an empirical hypothesis. The result-dependence hypothesis can also explain this phenomenon. If the results of an operation are a part of the definition of that operation, one need not make the same inference several times to make sure that it is correct. The definition of modus ponens includes the fact that it is ψ that follows from ϕ and $\phi \rightarrow \psi$. Thus, I do not need to infer ψ several times to make sure modus ponens is working. Cases where multiple proofs are required can be explained by the fact that they all involve proofs that are not surveyable or where the methods are contentious. In the former case, there is an issue as to whether one is in fact using a genuine mathematical, and hence result-dependent, operation. In the latter case, there is an issue as to whether one has actually engaged in the result-dependent operation involved. As we shall see when we come to the idea that proofs are memorable pictures, surveyability is a key component in result dependence.

Although the empiricist and result-dependence explanations are neck and neck on the first two themes, there are more things that result dependence can explain. Result dependence works especially well as an explanation of the role of imagination in mathematics. Imagination and result-dependent operations have an important property in common: in each case there can be no surprises. In the imagination, seeming is being. I cannot think that I am imagining something, but be wrong. This makes the imagination a lousy place for empirical experiments. Real experiments depend on one's ability to be surprised; one has to be able to get unexpected results. But the role of the imagination in mathematics is perfectly acceptable if mathematical operations are result dependent.

Result dependence implies that once one has conceived the operation, one has already conceived the answer, even if one is not fully aware of it. (This is why Wittgenstein says that surprise plays a different rule in mathematics than it does in empirical science. One is, as it were, surprised at oneself when one gets an unusual mathematical result, not at the world.) Thus it is perfectly possible for result-dependent operations to be performed entirely mentally. Therefore the role of the imagination in mathematics is incompatible with mathematics being experimental, but perfectly compatible with mathematics being result dependent. The relationship between result dependence and the image of proof as “unfolding” an “internal property” of numbers is also quite tight. An internal property of an object or process is a property that is necessary for the existence of the object or process. Thus if mathematical calculations are result dependent, the result is an internal property of the calculation. Therefore we feel like we are revealing an internal property because we are in fact revealing an internal property, not of the objects, but of the rule. The impossibility of widespread and consistent error in mathematical calculation is a logical consequence of the idea that mathematical functions are result dependent, combined with some tacit principle of limited comprehension. If functions are result dependent, then if we consistently calculated wrong, we would not in fact be calculating the old function. If we add a tacit principle of comprehension, we can go on to say there is a new function that we are in fact calculating when we consistently calculate an old function incorrectly. So again, the idea of result dependence helps explain a phenomenon of mathematical practice Wittgenstein was concerned with. At this point, result

dependence seems like a much stronger explanation of mathematical practice than the claim that mathematical statements are extremely well confirmed empirical statements.

Now, it is crucial to distinguish the claim that mathematical operations are result dependent from the claim made in the *Philosophical Remarks* and *Philosophical Grammar* that the meaning of a theorem is its proof. Both theses secure the same goal for Wittgenstein: they show that mathematical statements cannot be empirical statements by showing that one cannot get unexpected experimental results in mathematics. However, the claim that mathematical operations are result dependent is much more subtle and sophisticated than the earlier claims of the *Philosophical Remarks* and *Philosophical Grammar*. I see three main differences. First, the new thesis is a claim about mathematical *operations* not theorems. This difference is important because it moves the issue to a more fundamental level, where the nature of the rules in question is more apparent. It is simply more plausible to claim that it is a part of the definition of one-to-one correspondence that one must be able to link a hand and a pentacle in such a fashion than it is to claim that the definition of “there are infinitely many primes” is Euclid’s proof of that theorem. Second, Wittgenstein’s new thesis asserts that the result of a mathematical operation is a *part* of the definition of the operation, not the whole of it. This difference is important because it was the claim that the definition of a theorem is identical with its proof that forced Wittgenstein to say that one never proves the theorem one sets out to prove. By merely saying that the result is a part of the definition of the operation, Wittgenstein is leaving open the possibility that other associations, links to other language games and parts of culture, may form an essential part of the operation.

Finally, the thesis about result dependence is motivated by observations about mathematical practice, whereas the old thesis that the meaning of a theorem was its proof was motivated alternately by rank verificationism and a conception of the language game of proving as radically isolated from other language games.

Many of these same features serve to distinguish result dependence from analyticity. The two concepts are similar. Indeed, I think that the intuition which made philosophers want to say that mathematical statements were analytic was actually an undeveloped awareness of result dependence. Result dependence comes with less baggage, however. Because result dependence deals with operations and not propositions, we do not need to worry about truth by convention, truth by definition, or any of the other notions Quine has taught us to distrust. It also means that we do not have to worry about all of the formal problems of logicism, since we are not trying to construct the truth of all mathematics out of the truth of a few propositions.

The thesis of result dependence is quite far from the claim that the meaning of a theorem is its proof, or the claim that mathematical truths are analytic. It does, however, accomplish the same end as those earlier two ideas: it shows that mathematical statements are not empirical. (This is the first major antiempiricist thesis, and the seventh theme of Book 1 of the *Remarks on the Foundations of Mathematics*.) While a mathematical operation may have many connections with other language games, the outcome of that operation is built into it. Therefore, nothing in the empirical world forces the conclusion of a mathematical calculation on us. In this sense mathematics is all rules—the rules for manipulating the symbols are necessary and sufficient for

determining the result of the manipulation. But if the empirical world cannot affect the outcome of mathematical operations, mathematical statements can have no empirical content. Any empirical claim a mathematical statement might make would be unverifiable. Therefore, the claim that mathematical statements are not descriptive follows immediately from the result dependence of mathematical operations.

This leads one to wonder, though, about the purpose of mathematical statements. If they don't have any empirical content, what are they for? This is where our final theme comes in. Wittgenstein approaches the claim that mathematical statements are expressions of norms through the image of a proof as a memorable picture. At the core of the theme of proof as a memorable picture is the idea that a proof “is a *single* pattern, at one end of which are certain sentences, and at the other end a sentence (which we call the ‘proved proposition’)” (*RFM I* §28). On the surface, this statement is completely trivial—obviously proof and proposition proved must form a whole—but what Wittgenstein means by it is that the proof and the proposition proved are *used* as a single unit. The exact nature of this use is unclear. However, it is definitely linked to the fact that mathematical statements have no empirical content. Wittgenstein writes that the “experimental character” of a sequence of activities “disappears when one looks at the process simply as a memorable picture,” (§80) and, “the proof does not serve as an experiment; but it does serve as the picture of an experiment” (§36). Wittgenstein has a strong tendency to say that the proof and proposition proved are used as a single unit by forming a paradigm for future actions. For instance, Wittgenstein says that if we view the diagram of the hand and the pentacle as a proof, we will use the picture of the two figures

and the connections between them to judge future correlations of groups of objects. As a result, any two groups we decide to correlate in the future will not be compared with each other, but with the figure of the hand and pentacle. Describing the two sets of objects in a future comparison, Wittgenstein says, “we do *not* correlate them, but *instead* compare the groups with those of the proof (in which indeed two groups are correlated with one another.)” (§31). This is a way of viewing the hand and pentacle as expressing a norm. The proof demonstrates what is to count as a good correlation. Thus, Wittgenstein writes, “The proof doesn’t *explore* the essence of the two figures, but it does express what I am going to count as belonging to the essence of the figures from now on” (§32). Remarks 79 and 80, where he actually introduces the phrase “memorable picture,” imply a similar view. In those passages he asks us to consider the possible positions of a puppet. The display of such positions could convince one of a possible way of comporting it. In such a case, we would only be interested in the positions that were memorable and easy to reproduce. It is in this context that Wittgenstein says that the experimental character of a sequence of actions disappears when one considers it as a memorable picture. The implication is that the picture of a puppet in a specific position is to be used as a guide in moving similar puppets.

Viewing a proof as a guide to future actions is one way of viewing it as an expression of norms. It makes proofs analogous to promises or New Year’s resolutions. They say, “From now on, I will do things that look like this.” There are problems with this kind of normativity, however. The examples in the last paragraph deal with simplified examples Wittgenstein believes are analogous to proofs, rather than with actual proofs, and the first

problem that comes up with this way of understanding the proof as memorable picture theme is understanding how it might apply to real proofs. In a constructive proof, the proof can function as a model for future applications of a rule. Euclid's proof of the infinity of primes, to take an example that Wittgenstein often discusses, shows one how to find a prime larger than any given prime. The pattern of finding a prime in the proof thus serves as a model for how to find primes in the future. But not all proofs are constructive proofs, and it is hard to see how existence proofs and proofs that a certain proof or construction is impossible can be made to fit Wittgenstein's conception of proof. But this is the least of the problems for this way of conceiving of a proof as a memorable picture or expression of norms. As I pointed out during the hand-and-pentacle discussion, the use of proofs Wittgenstein describes is directly at odds with his discussion of rule following. We seem to be asked to use proofs the same way the image of a cube was used as a definition of the word cube. One might try to argue that the situations are not really analogous because in the cube example we were taking the cube to be constitutive of the definition of the word cube, whereas when we take a proof as a picture we are simply talking about a technique that would be consciously applied. Wittgenstein never was suggesting that it is impossible to compare an image of a cube to an actual cube. He was merely saying that this was not constitutive of the meaning of the word "cube." Similarly, the defense of Wittgenstein might say, there is nothing wrong with using one correlation to judge future correlations, so long as one does not take this as constitutive of the meaning of "correlation." Unfortunately, the meaning of "correlation" is precisely what is at stake here. When talking about proofs as memorable pictures, Wittgenstein says things

like “the mathematician creates *essences*” (§32). When one produces a proof, one creates a new definition, which is deposited “among the paradigms of language.” A proof cannot be both an essence and a picture that serves as a model for future action.

There is, however, another way of understanding proof as a memorable picture. Rather than guiding future applications of a rule, a proof might serve to state explicitly a rule that may have otherwise only been implicit in one’s actions. In making this suggestion, I am inspired by Brandom’s approach to logic in *Making It Explicit* (1994). Brandom claims that normativity exists primarily as an implicit aspect of our practices. When someone makes an utterance, say about what they believe, other people treat them as committed to certain courses of action. When a group of people evaluate each other’s behavior in this manner, a norm is implicit in their behavior. According to Brandom, the purpose of logical vocabulary, e.g., the sentential connectives, is to generate explicit statements of these norms. For instance, the conditional allows us to state explicitly the idea that when one person has one commitment they thereby have another. The account I am suggesting here extends this thesis to mathematical vocabulary. The purpose of mathematics is to render explicit a set of norms implicit in ordinary practice. The norms that mathematical statements make explicit are precisely the norms found in the sorts of actions Wittgenstein examines. For instance, when we place two pairs of apples on a table and find that there are only three apples, we are committed to looking for the missing apple, or at least believing that an apple has disappeared. This commitment, and many other commitments like it, is expressed by the claim that $2 + 2 = 4$. Part of this new way of viewing mathematical statements as expressions of norms is a shift in what counts

as the basic unit of normative expression. For Wittgenstein, the smallest utterance or inscription that expressed a norm was a mathematical proof. It was the proof that was said to be a memorable picture. I would like to say instead that the proposition is like a memorable picture. The reason it is memorable is not so that it can guide future action, but so that it can capture what many otherwise heterogeneous actions have in common. The proposition, I think, is a much more natural unit of expression than the proof. We do not need to worry about how a long proof can guide or summarize action, and we can return proofs to their traditional role of convincing us to accept propositions.

The Historical Objection

Even in this modified form, the Wittgensteinian view of mathematics is radical, and the reader no doubt has many objections. Right now, however, I will only deal with objections that might be raised by a mathematical empiricist like Philip Kitcher. More fundamental objections will be dealt with in the last chapter. The first response I will look at here is a plea for compromise. Kitcher, as we saw in the last chapter, believes that all the empirical input into mathematical knowledge occurred at the beginning. Perhaps we could say that for a period of time mathematical statements were descriptive, but once they attained a certain degree of confirmation, the descriptive aspect dropped away, and they became statements of norms. We could go so far as to specify a time when this occurred in various mathematical traditions. There is a point in the history of written numbers in Mesopotamia where number symbols shifted from being terms for specific kinds of objects—flocks of sheep, where a flock was understood to specifically be 10

sheep—to generic numbers which could be used in combination with any object symbol (Schmandt-Besserat 1992, 191; 1996, 117). Perhaps this could be seen as a sign that number had become a pure system of norms. If we find that we can actually gather enough historical data to decide this issue, we might even go on to look for other sorts of knowledge that have evolved according to this pattern. Wittgenstein is certainly amenable to this kind of thinking. Twice during the *Remarks on the Foundations of Mathematics* he suggests that mathematical statements are “empirical statements hardened into a rule” (VI §§22, 23). In the same place he refers to “withdrawing a proposition from being decided by experience.” In his last work, *On Certainty*, he makes the same point more generally, comparing human knowledge to a river with the hard propositions forming a bed for the fluid propositions. “It might be imagined that some propositions, of the form of empirical propositions, were hardened and functioned as channels for such empirical propositions as were not hardened but fluid; and this relation altered with time, in that fluid propositions hardened, and hard ones became fluid.” (OC §96). Wittgenstein adds, “And the bank of the river consists partly of hard rock, subject to no alteration or only to an imperceptible one, partly of sand, which now in one place now in another gets washed away, or deposited” (OC §99). So here we have an easy compromise for Kitcher and Wittgenstein: mathematical propositions are like sedimentary deposits along a river bank. They once were a part of the rushing river, now they guide it.

Unfortunately, I don’t think this compromise is compatible with the core views of either Kitcher or the Wittgenstein of the *Remarks on the Foundations of Mathematics*. For Kitcher, there is little point in calling himself an empiricist if the empirical content of

mathematics has vanished. Kitcher is quite aware of this, and has gone out of his way to develop a semantics for mathematical language that will allow it to retain its empirical content despite continual shifts in use and metamathematical understanding. He wants to be able to say, for instance, that when the mathematicians of the nineteenth century formalized our concepts of irrationals, they “enabled us to specify the referents of expressions which had been in use since antiquity” (1984, 126). Drawing on the work of Saul Kripke (1972) and others, he points out that one can successfully refer to an object without knowing enough about it to uniquely identify it. One can refer to Einstein without knowing anything that would distinguish him from any other famous scientist. According to the line offered by Kripke (and his predecessors and successors) reference is possible in such situations because the referent was fixed in some initial baptismal ceremony, when the meaning was established by ostension or description. Kitcher accepts this description of reference, but adds the idea that the link between word and object can be renewed periodically. This allows a term to acquire a diverse set of possible referents, which Kitcher dubs the *reference potential*. Because a reference potential can shift and grow, we can initially refer to an object without having any correct descriptions of it. Later, when correct descriptions are introduced, the terms acquire a heterogeneous reference potential, which allows the proponents of the new theory to convince the proponents of the old that the objects they have been referring to are really best described in the new way. Setting aside the issue of whether this concept of reference works or not, it is clear that Kitcher is committed to the idea that mathematical language has always

had the same referents, even if they have not always been known. Therefore it would be impossible for it to change status from descriptive to normative.

On Wittgenstein's side, there is nothing to prevent the arguments he uses to show that mathematics is currently normative from being applied to any era of history. While it is probably true that statements can change their status from normative to descriptive and back again,¹⁵ Wittgenstein has proved a very strong thesis about mathematics: it can only be normative and still be mathematics. Wittgenstein showed that mathematical language was purely normative by showing that mathematical operations were result dependent. Could there have been a time when mathematical operations were not result dependent? Well, there's a cheap way to say that they could not have been: Wittgenstein has shown result dependence is a part of the definition of mathematics, so any past activity that did not involve result-dependent operations could not, by definition, be mathematics. I do not want to take so facile an approach, so let me ask the question this way: could there have ever been an activity that is like contemporary mathematics save for the fact that it does not involve result-dependent operations? Let's perform a modestly historically informed thought experiment using the accountants in the ancient Mesopotamian city of Uruk discussed previously. Suppose it is the job of one such accountant to record on one side

¹⁵ It is a mistake to characterize the contrast between norms and descriptions simply by saying that descriptions are revised and norms are not. In the passages from *On Certainty* and *Remarks on the Foundations of Mathematics* where Wittgenstein discusses hardening empirical propositions into a rule, he assumes that this is done by making the proposition immune from the influence of experience. This assumption could be right or wrong, depending on what you think it means. If it means empirical refutation it is perfectly correct. However, norms do change over time, and statements about the world do play a role in the debates over how norms should change.

of a tablet the size of several different offerings of grain to the temple, and to total these offerings on the reverse side. The accountant could use many means to find this total. He could collect tokens representing the units of grain in each offering and then count the total. He could use an abacus or counting board. He could count upwards, starting from the number of the first collection, keeping track of the number of numbers he as recited, stopping when the number of numbers is equal to the second collection. This process would be repeated until he had totaled all the offerings. He could even consult a table of sums. Let's assume, though, that no one has ever summed numbers this large before. (It's a banner year for the temple and offerings are unprecedentedly large.) This would rule out consulting a table to get the answer. I maintain that all of the other operations our ancient accountant might use are result dependent. I believe this because the operations his adding methods are composed of are all result dependent. Counting up from the number of one collection a number equal to the number of another collection involves two activities, counting, and coordinating different acts of counting. As we shall see in the next chapter, counting involves the ability to put objects in one-to-one correspondence, the ability to bear in mind a stable sequence of number words, and the convention that the last number word uttered represents the cardinality of the set counted.¹⁶ All of these are result-dependent operations. The act of putting collections in one-to-one correspondence was shown to be result dependent by Wittgensteinian thought experiment like the hand-and-pentacle example. Similarly, it is a part of the definition of

¹⁶ This analysis of counting comes from Gelman and Gallistel (1978), whose work is part of the subject of Chapter 3.

keeping a stable sequence in one's mind that the sequence one ends with is the same as the sequence one begins with. Coordinating two different acts of counting is essentially a matter of putting ordered sets in one-to-one correspondence, yet another result-dependent operation. The technique of collecting piles of tokens corresponding to the offerings of grain and then counting their union is also composed of result-dependent operations. Creating piles that correspond to the offerings of grain is a matter of putting sets into one-to-one correspondence, which we know to be a result-dependent operation. We also know that counting the union of these collections is a result-dependent operation. All that remains is the operation of actually forming the union. I believe that this is also a result-dependent activity, even in situations like this one, where one has no preconceptions about what the result of the union should be. There are two basic ways something could go wrong here. The obvious one is that an object could be misplaced. This sort of mistake can be checked for empirically—we just look around and see if a token hasn't rolled under our desk. On the other hand, uniting two collections could yield a wrong result if objects in those collections meld into each other. (Think of trying to count two collections of water droplets.) To rule out this kind of mistake, we must rely indirectly on the result of summing the collections. To know that two objects haven't merged into each other when we united the sets, we must rely on our concept of an individual object. The correct sum of the two collections is implicated in this concept. It is a part of the definition of an individual object that it is not another object. Therefore it is a part of the individuality of two objects that they form a collection of two. Similarly, the separateness of three objects guarantees that they form a collection of three, and so on. Therefore, in

saying that the individual objects remain separate in uniting the two collections we have implicitly fixed the size of the result.¹⁷ Operations involving counting boards and abacuses would work the same way as collecting tokens, with the possible addition of place value techniques, which would also clearly be result dependent. Finally, you should note that if any of these methods are result dependent, they all must be, because their results must match. The sum of two numbers cannot be a norm if computed one way and a description if computed another.

This all might be quite puzzling. How can a summation be result dependent if we do not know what the sum is? How can the ancient accountant ever be wrong? How can he check himself? These are all “how was it possible” questions. The answer is, “the same way that it is possible now.” Even now large sums are computed by people who do not know the answer in advance, yet we can say that their actions are result dependent. One does not need to know the result of a result-dependent operation in order to perform the operation. All that is necessary is that the correct result be used as a criterion for judging whether the operation was performed correctly. The situation was no different thousands of years ago in Mesopotamia. In fact, at no point in the preceding thought experiment have I made real use of the historical setting. This is because the considerations Wittgenstein raises about mathematics are ahistorical. The key is the idea of result dependence. If there were some other factor that led us to say that mathematics was purely normative, there might have been a chance that the status of mathematics could

¹⁷ This argument is quite sketchy. I will be able to give a more formal argument in Chapter 5.

have changed. But there is nothing about the result dependence of mathematical functions that is tied to a particular era.

The Application Problem

If mathematics is purely normative, how come it's so good at yielding empirically accurate results? Explaining the ever more astounding achievements of mathematical science is a perennial problem for any theory of mathematical knowledge. Indeed, many people seem devoted to making the success of mathematical science inexplicable. The stated goal of Eugene Wigner's famous essay "The Unreasonable Effectiveness of Mathematics in the Natural Sciences" is to show that "the enormous usefulness of mathematics in the natural sciences is something bordering on the mysterious and that there is no rational explanation for it" (1960). The advantage of an empiricist theory like Kitcher's is that it is able to explain the usefulness of mathematics. The position I have advocated, on the other hand, has no resources to answer this problem. It appears as though mathematical science is the product of the most bizarre of coincidences: although mathematics is in no way a description of the world, it happens to be amazingly effective at helping us design bridges. If one believes that mathematical statements are normative and not descriptive, there is only one response to this "application problem." In classic Wittgensteinian fashion, it must be dissolved rather than solved. We have to realize that there is nothing to be explained here. The phenomena we are worried about—the success of mathematics in the sciences—is a product of a particular philosophic view of what

mathematics is. Once we completely rid ourselves of this view, we will stop asking for explanations.

Really, the application problem, when phrased as an objection to a philosophic theory of mathematics, is a species of so-called “inference to the best explanation” arguments for scientific realism. Frequently it is argued that the only explanation for the success of natural science is that its statements are true (Boyd 1984 is the best attempt at this sort of argument). Given the principle of inference that says that the propositions that best explain the world around us are true, it then follows that the proposition “the statements of natural science are true” is true. Similarly, an empiricist could argue that the best explanation for the success of mathematics in the sciences is that mathematical statements are true descriptions of aspects of the world (such as “structural features in virtue of which we can segregate and recombine objects”). Given the principle of inference to the truth of the best explanation, we can then infer the truth of mathematical empiricism. However, at this point there are a host of replies to this sort of argument. In the elegant essay “The Truth Doesn’t Explain Much” (1980/1983) Nancy Cartwright argues that the explanatory function and the truth-telling function of science are separate. Thus science could give successful explanations of the phenomena and still not be true. A host of others have made similar arguments, either attacking the ground level assumption that the truth of science could explain its success, or attacking the principle of inference applied to that assumption, the principle that says we should then infer the truth of the statement “scientific statements are true.” (Van Fraassen 1980, 97–101; Duhem 1914/1991). The most cutting argument comes from Arthur Fine (1984/1986), who has

pointed out a basic circularity in inference to the best explanation arguments. To show that our good scientific explanations are true, it invokes the idea that good explanations must be true.

All of these arguments could apply to the application problem, but they would not do the job a real Wittgensteinian dissolution would. Although they argue that those who would use the application problem as an argument for empiricism commit a fallacy, they do not show us how language has bewitched us into thinking that there is such a thing as the application problem at all. The origin of this mistake is a false analogy between judging the usefulness of a particular mathematical theory in explaining a particular phenomena, and judging the usefulness of mathematics in general in explaining the world in general. The former sort of judgment is made all the time, but the latter judgment cannot sensibly be made. When have the controlled clinical exams been performed comparing mathematical science to nonmathematical science? We always compare different equations, never an equation to a lack of an equation. The truth is, mathematics forms the language tests are framed in. This becomes obvious when we try to imagine what a test of the reliability of mathematical science would look like. Would it be quantified? We can compare the accuracy of Ptolemaic and relativistic astronomy, because they both offer predicted measurements for the positions of heavenly bodies. But what task could we compare relativistic astronomy and a hypothetical nonmathematical astronomy in? When we switch from talking about individual theories to talking about mathematics as a whole, we switch from talking about a linguistic structure whose job it is to describe the world to one whose job it is to set the rules for how one describes the

world. We are fooled by the fact that mathematics appears in all our scientific theories of the world into thinking that it is the mathematics that makes them powerful. (The impression is no doubt fueled by the fact that the mathematics is often the hardest part of a theory to understand.) But this is like noticing that whenever a person swims really fast, they are always swimming in a liquid, and inferring that liquids makes one swim well. Really, saying that mathematics is good for science is like saying that liquids are good for swimming in.

In “The Unreasonable Effectiveness of Mathematics in the Natural Science” one can see Wigner falling into just this very trap. Wigner describes two roles for mathematics in science: it is used to formulate the laws of nature and to check those laws against experimental results. One might think that this would lead him to conclude that math is not an effective tool, but a framework for developing and evaluating tools. But instead Wigner declares that the most important function of mathematics is its role in formulating scientific laws because it is there that math is “sovereign” (1960, 6). Once this function is isolated, math takes on mysterious powers. How is it that mathematical concepts and laws, designed primarily to show off the intellect of the mathematician, could be so good for formulating laws of physics? As soon as he puts the question this way, Wigner falls into aporia: “It is difficult to avoid the impression that a miracle confronts us here.” But it is no more a miracle that we can formulate accurate laws mathematically than it is that we can test them mathematically. The two operations form a closed system. When we look at one part of mathematics in isolation it seems like an incredible application, but when we look at the whole of it, it ceases to look like an application at all.

Wigner actually has to do a fair amount of work to maintain the illusion that something miraculous is going on here. He notes that only a small portion of the mathematical concepts ever developed are used in physics, and that very often the physicists develop these concepts in the course of their work independently of mathematicians (ibid., 7). However, he dismisses these points, saying only that they do not make the successful application of some mathematical concepts any less unlikely. He admits that mathematical language is “the only language we can speak” in natural science, yet he quickly adds that mathematics also “is, in a very real sense, the correct language” to describe the world. (ibid., 8). But if Wittgenstein has taught us anything it is that it is only possible to speak of correctness where it is possible to speak of falsity. If mathematics is the only language science can use, it is pointless to say it is “correct.” There is no incorrect language to contrast it with. Really, to create the impression that there is a mystery here, Wigner has to set aside a great deal of common sense.

Strangely, although Wittgenstein has all the techniques he needs to dissolve the application problem, his treatments of the issue are very unsatisfactory. The remarks that are most relevant to the application problem are a sequence of comments that appears in manuscripts as early as 1931. The sequence turns up again in a more polished form in the *Philosophical Grammar*, and again, slightly altered, in the *Philosophical Investigations*. The sequence of remarks is fairly ambiguous, but on at least one reading, they are a complete mishandling of the issue. The comments appear in their baldest form in a manuscript Wittgenstein showed to Friedrich Waismann in September of 1931

Why do men think, why for instance do they calculate the dimensions of a boiler and not rather leave it to chance what size will come out? Will

this calculation perhaps save us from an explosion of the boiler? No, the boiler can explode despite the calculation. But men will no more dispense with calculating the dimensions of boilers than they will put their hands into a fire once they have been burnt

...

Now if I am asked, Had you any right to make the boiler 15 mm thick? Can you sleep soundly? Then I cannot help replying with a counter-question: What does 'right' mean here? If what you mean by it is that we know that an explosion of the boiler is impossible, then I had no such right. But if by 'right' is meant that I have calculated the dimensions of the engine in terms of this calculus, then I do have the right. There is no more that can be said. (*WWK*, p.171).

Here Wittgenstein seems to be denying the obvious truth that we have very good techniques for designing boilers. Worse, his reason for doing so is that we cannot be apodeictically certain that the boiler won't explode: "Will this calculation perhaps save us from an explosion of the boiler? No, the boiler can explode despite the calculation." In the second portion of the quotation, he seems to justify this approach by saying that one can only judge whether one has done something correctly in the context of a calculus. Thus we can say we built the boiler right from the perspective of a particular boiler-building technique, but we cannot say in general that we have built the boiler well. If Wittgenstein thinks that we cannot justify our boiler-making techniques, why does he think we use them? In this passage he implies that we calculate as a sort of reflex generated by our fear of calamity. But the idea that we cannot say in general whether we have built the boiler well flies in the face of common sense. Although Wittgenstein claims that there is nothing more to be said here, there is really plenty more. We can compare the statistical effectiveness of various boiler-making techniques; we can even

throw in the effectiveness of boilers whose dimensions were generated at random as a kind of control. This is simply common sense.

Different aspects of these ideas are developed in the version of this passage that appears in the *Philosophical Grammar*.

What does man think for? What use is it? Why does he *calculate* the thickness of the walls of a boiler and not leave it to chance or whim to decide? After all it is mere fact of experience that boilers do not explode so often if made according to calculations. But just as having once been burnt he would do anything rather than put his hand into the fire, so he would do anything rather than not calculate for a boiler.—Since we are not interested in causes, we might say: human beings do in fact think: This for instance is how they proceed when they make a boiler—Now, can't a boiler produced in this way explode? Certainly it can.

...

What the thought of the uniformity of nature amounts to can perhaps be seen most clearly when we fear the event we expect. Nothing could induce me to put my hand into a flame—although after all it is *only in the past* that I have burnt myself.

The belief that fire will burn me is of the same nature as the fear that it will burn me.

Here I also see what “it is certain” means.

...

“But after all you do believe that more boilers would explode if people did not calculate when making boilers!” Yes, I believe it,—but what does that mean? Does it follow that there will in fact be fewer explosions?—Then what is the foundation of this belief? (*PG I §67*)

Fortunately, in this passage Wittgenstein also backs away somewhat from his more abrupt dismissal of the power of modern engineering. At the place where he previously remarked, “No, the boiler can explode despite the calculation,” he now says, “After all it is a mere fact of experience that boilers do not explode so often if made according to calculations” Wittgenstein has also dropped his hard stance on the intelligibility of saying that we have in general built the boiler correctly. In place of that discussion we now have

a colloquy between Wittgenstein and his imaginary friend where he claims to believe that we use the best boiler-making techniques, but not to know why. What is most interesting about this passage, though, is the way he expands on the idea that we calculate out of reflex. Our very belief that the future will be like the past, our attempts to plan for the future on the basis of the past, are associated with the emotional reaction we had to past events, as if planing were a sort of cringing. Most of the passage in the *Philosophical Grammar* reappears in the *Philosophical Investigations*. The quotation above up to the first ellipsis forms §466. The section up to the second ellipsis, minus the last line about certainty becomes §472–73. Between these remarks, he now has comments that deal with the reasons for thinking in general. A representative quote is “Does a man think then, because he has found that thinking pays?—Because he thinks it is advantageous to think?” On the whole the tone is the same as it was in the *Philosophical Grammar*.

The versions of these comments found in the *PG* and *PI* remain problematic, I think, because they flit back and forth between examples that are too specific and examples that are too general. In these passages he is still hinting at the idea found in the earliest remarks that we cannot really say our boiler-making techniques are effective. But at this very specific level, we can say our techniques are effective. Their specificity means that there is a normative context that makes comparison possible. On the other hand, he also wants to talk about the effectiveness of thinking in general—he tries to answer the question he originally put to himself, “What does man think for?” But here again I think it is possible to say that activity in question accomplishes all sorts of goals. One can say, for instance, that an organism that thinks is better suited to the evolutionary niche we

occupy than one that doesn't. At this level of generality, we can consider all sorts of goals, not just accurate description of the world. So again we have plenty of normative contexts to judge the effectiveness of our techniques. Neither the specific remark about boilers nor the general remark about thinking strikes its intended mark, because none of them address the issue of the application of *mathematics* to the world. Mathematics cannot be said to be successfully applied to the world because there is no context where it is not applied that we can compare its application to. This in turn is true because mathematics is a basic part of the normative context that allows us to compare descriptions of the world. Wittgenstein's goal in these remarks seems to be to discuss the application problem. The design of boilers or bridges or other engineering-intensive artifacts we risk our lives on are classic examples of the application of mathematics to the world. They are the stuff the application problem is made of. But Wittgenstein doesn't quite seem to hit on what I think makes such examples misleading, that is, the difference between talking about the application of this or that mathematical formula, and the application of mathematics in general.

Conclusion

In this chapter, I have attempted to identify and refine Wittgenstein's strongest argument against mathematical empiricism. The argument of his middle period works, the *Philosophical Remarks* and the *Philosophical Grammar*, was examined and rejected. Instead I have advocated an argument I find in his later works, the *Philosophical Investigations* and the *Remarks on the Foundations of Mathematics*. The argument comes

in two parts. The first, found largely in the *Investigations*, claims that normativity is irreducible. The second part of the argument was found in Book I of the *Remarks on the Foundations of Mathematics*. I examined several motifs from this seemingly desultory work, and singled out two as theses whose support comes from the others by a roughly abductive argument. The first of these is the idea that mathematical operations are result dependent. As a result, the rules that govern the use of mathematical statements are not only irreducible, they are all there is to math: They completely determine the result of mathematical operations. The second thesis is the idea that mathematical statements are used to explicitly state normative commitments that would otherwise be implicit. This claim was then defended against two possible objections from empiricists.

There are, however, other alternatives to empiricism besides the thesis that mathematical statements serve to express norms. Perhaps the evidence presented only shows that mathematics is a priori or innate? To see whether this is true, I would like to spend Chapter 3 developing a plausible nativist and apriorist position, just as I did in Chapter 1 for empiricism. At first the nativist and apriorist position will be developed purely in opposition to empiricism. In Chapter 4 I will begin to compare it with the Wittgensteinian considerations I have brought forth here.

Nativism and Apriorism

Introduction

The argument of the last chapter adduced several aspects of mathematical practice as evidence that (1) mathematical operations are result dependent and thus mathematical statements cannot be descriptions of the world and (2) mathematical statements are actually expressions of norms. However, rather than demonstrating these positive theses, the argument of the last chapter might actually only show that mathematical empiricism is false. The mathematical empiricist, remember, argues that mathematical knowledge is *learned from* experience and *justified by* experience. To deny that mathematical knowledge is learned from experience is to espouse a form of nativism.¹⁸ The nativist claims that some kind of knowledge, in this case mathematical knowledge, is a product of structures of the brain that are present in us from birth. To deny that mathematical knowledge is justified by experience is to espouse a form of apriorism. The apriorist claims that for some kind of knowledge, in this case mathematical knowledge, it is possible to justify the belief without reference to experience.

The previous chapter may have only shown that empiricism is false because the nativist and the apriorist can do as good a job as the Wittgensteinian in explaining the

¹⁸ This view sometimes gets called “rationalism” or “innatism,” especially by those writing about the history of philosophy.

aspects of mathematical practice adduced as data in the last chapter. So although the hypotheses of result dependence and the normativity of mathematical statements enjoy more abductive support than mathematical empiricism, they do not edge out the nativist and the apriorist. Consider how the nativist would deal with the aspects of practice adduced in the last chapter, beginning with our habit of discrediting experiences that do not conform to our mathematical ideas, rather than using those experiences to discredit the mathematics. Wittgenstein noted that if one puts two pairs of two apples on a table and finds that one only has three apples in sum, one says that one of the apples has disappeared, rather than saying that two and two do not equal four. The mathematical nativist has a very easy time explaining this phenomenon. We are hardwired to think that two and two are four, so we must question our experience, rather than our belief. In fact, the nativist can claim that she has both a more precise account of the phenomenon Wittgenstein is describing and a more detailed explanation of that phenomenon. As we shall shortly see, a great deal of the empirical research into mathematical nativism is dedicated to showing that at a very young age children, and even infants, will express something like surprise when presented with situations like Wittgenstein's thought experiment with the apples. Thus rather than simply having a thought experiment, the nativist actually has hands-on experience with the situation Wittgenstein describes. The nativist would also claim that her explanation of such phenomena is more detailed than Wittgenstein's. Rather than vague remarks about mathematics playing a normative role, the nativist has an empirical hypothesis. As we shall see, many believe that an infant

possesses an innate representation of facts like $2 + 2 = 4$, and the mismatch between this representation and the phenomenon motivates infant behavior.

Of course, our unwillingness to revise mathematical beliefs was something that even the mathematical empiricist could account for. It took several other pieces of evidence to argue for result dependence. However, the nativist can account for these as well. Wittgenstein thought that mathematics has a special relationship to the imagination because we can calculate purely in the realm of thought. But this would be expected if there were innate representations of mathematical facts. We can calculate purely in the medium of the imagination because we can access our innate mathematical representations and manipulate them. Wittgenstein claimed that mathematical proofs had to involve types of objects or kinds of objects, rather than particulars. The nativist could say that this is because we calculate with the symbols in our mind, which can stand for an infinite number of actual collections. Wittgenstein says we cannot imagine consistently performing a mathematical operation incorrectly, because if we did, we would simply be performing another operation. The nativist can say that this is because the only operations we can imagine performing consistently are the ones we have hardwired into our minds. If we are not performing one of them, we are performing another. Wittgenstein makes much of the idea that when we calculate, we believe we are “unfolding” a property already in the collection. The nativist can reply that the place where these properties preexisted is the mind.

The apriorist can also easily explain the Wittgensteinian phenomena of Chapter 2. The apriorist believes that we have some kind of nonexperiential access to mathematical

facts, such as an intuition. The strength of this intuition could certainly explain our habit of discrediting experiences that do not conform to our mathematical ideas. It would also easily explain the tie between mathematics and the imagination, the sense we have that when we perform a calculation, we are unfolding a pre-existing property, and the generality of mathematical knowledge. Mathematics is performed in the imagination because we have nonexperiential access to mathematical facts. We have the sense that mathematics unfolds essences because the mathematical facts we have access to are essences, or alternatively are necessary structures of the human mind. The same special status of the facts we have nonexperiential access to can explain why mathematics always takes place at the level of types. Not only are these sorts of things explicable by the apriorist, they have traditionally been used in arguments for the existence of an a priori intuition. The impossibility of widespread, consistent error in mathematics could also be explained in terms of the power of our mathematical intuition. If the intuition serves as a limit on what we can conceive, then in order for someone to be calculating a function consistently, but incorrectly, they must be calculating another function.

So we see that nothing that has been said so far rules out nativism or apriorism. We are thus brought to the second wing of the dialectic of this dissertation. The purpose of this chapter is to develop plausible versions of both the nativist and the apriorist positions, just as the first chapter developed a plausible empiricism. Rather than providing a unified account, this chapter will pursue three different tracks. One will develop nativism as a free-standing position, one will develop apriorism as a free-standing position, and the last will combine the two. The difference between these views

comes in how they define the terms ‘innate’ and ‘a priori’. The first two positions adopt a very strong definition of either innateness and apriority and then argue that some aspect of mathematics fits that definition. The third position adopts very weak definitions of apriority and innateness—so weak that the two terms become co-extensive—and argues that aspects of mathematics fit these weak definitions. These different takes on nativism and apriorism will receive different treatments in Chapter 4. I will argue that the entities postulated by two stronger positions are unable to perform the functions assigned to them, and moreover are unlikely to exist on empirical grounds. The weaker position, by contrast, will simply be shown to be compatible with the Wittgensteinian account of mathematics. Indeed, in a certain way it is appealing to a neo-Wittgensteinian, because it points the way for an integration of empirical work into the Wittgensteinian account of mathematics.

The job of this chapter, however, is simply to clarify the various nativist and apriorist positions and to provide some evidence for the most plausible versions of them. My first step will be to clear away two definitions of innateness that are completely unworkable, despite their popularity in the literature. The first is based on a notion called “domain specificity.” This definition is extremely popular among cognitive developmental psychologists, largely, I think, because of the internal politics of developmental psychology, and not because of any intrinsic merit. The other definition I would like to dismiss is a definition based on the notion of causality. Although this is not quite as popular as the domain-specificity definition, it still appeals to some because of its simplicity. Once I have cleared away the definitions of innateness that simply don’t work,

I will develop the first full-fledged nativist position, based on the notion of an innate representation, outlining the claims of the position and the empirical evidence for these claims. The next step will be to outline the first full-fledged apriorist position, which will be based on the notion of an a priori intuition. This time the evidence for the position will naturally be rational, not empirical. Finally, I will propose a third position, which uses weaker notions of innateness and apriority, and claims that the two concepts are co-extensive. Because the position is both nativist and apriorist, it will be backed by both empirical and rational evidence.

Two Unworkable Definitions of 'Innate'

Nativist theories postulate the existence of one or more of a set of three objects: innate ideas, innate knowledge, or innate abilities. When I want to refer to these innate objects generically, I will refer to “innate capacities.” During the enlightenment, debate focused on the first innate capacities. John Locke opens his *Essay Concerning Human Understanding* by attacking the idea that humans innately possess “principles, both speculative and practical,” such as “It is impossible for the same thing to both be and not to be” or “that one should do only as one would be done unto” (1689/1964, 67, 79). Innate principles such as these are what I am calling innate knowledge. One of the ways Locke attacks the existence of such innate knowledge is by denying the existence of one of its components, innate ideas. If we innately know that it is impossible for a thing to both be and not be, then we should have an innate concept of being. But, Locke thought, this was impossible. To this pair of innate capacities the modern discussion adds a third:

innate abilities. Innate abilities are to be thought of as a species of Gilbert Ryle's *knowing how*. Ryle distinguishes the *knowing that* a proposition is true from *knowing how* to perform an action, a knowledge that one may not be able to put into words at all. The thing to bear in mind about knowing how is that it is also distinguished from mere correct action. A clock may keep good time, but it does not *know how* to keep time. Knowing how requires something more—that one be capable of governing one's actions, to somehow reflect and improve upon them (Ryle 1949, 28). If an ability is to be considered innate, it must involve an innate capacity to govern oneself in this fashion.

Nativist views, no matter what sort of innate capacity they champion, face the same basic challenge. Children are not born counting, or attesting their knowledge that two and two are four, or proclaiming their mastery of the number concept. Things have to happen to them before they show signs of having these capacities. Traditionally, advocates of innate capacities avoid this problem by saying that the capacity believed to be innate lies dormant until something "triggers" it. This talk of triggering is not entirely vague. Consider the case of innate knowledge. Intuitively, a process counts as triggering innate knowledge if the *content* of the knowledge gained can be seen as already present in the brain at birth. Conversely, if the content of the knowledge is more in the structure of the experience one has in the world, we say that the process is learning. It is not hard to see ideas and abilities as having contents that would lend themselves to similar accounts of innateness. The difficulty in each case is coming up with a functional understanding of what it is for content to lie either in brain structures or in the structure of experience.

Versions of nativism can easily be distinguished by looking at the mechanism they use to locate the content of knowledge in the brain or in the environmental stimulus.

The first version of nativism which I wish to dismiss in this subsection uses the notion of a *domain-specific constraint* to locate the content of knowledge. The phrase ‘domain-specific constraint’, and its correlate phrase, ‘domain-general constraint’, are terms of art in developmental psychology. The first psychologists who studied learning believed that knowledge acquisition worked basically the same way no matter what knowledge was being acquired. Jean Piaget, the founding father of developmental psychology, believed that all aspects of learning were governed by processes he termed *assimilation*, *accommodation*, and *equilibration*, activities of transforming new information to fit one’s existing way of thinking, adjusting one’s way of thinking to fit new data, and balancing the two. B. F. Skinner’s associative learning was similarly a theory based on a domain-general process. Researchers since the sixties, however, have been more likely to postulate separate areas of cognitive growth (language development, perceptual development, etc.), each governed by unique rules. These rules are thought of as *constraints* on the way the child can develop. They ensure that no matter what experiences a child will have, her mind will not take a certain form, or at least be very resistant to taking that form. For instance, one of the principles we will look at is the *stable-ordering* principle. It says that one must always use the number words in the same order while counting. If the principle is operating in a child’s mind, she will be disinclined to change the order of the count words, no matter what experiences she has in learning to count. Although the language of constraints is a language of bondage, the

effect is liberating. Constraints allow a child to develop very elaborate mental structures with very little input.

The notion of a domain-specific constraint is very appealing to developmental psychologists largely because it highlights the idea that knowledge arises from the interaction of innate capacities and the environment. Older discussion of innateness made innate knowledge an all or nothing affair. Either knowledge was acquired through a learning process, or it was present innately. Domain-specific constraints, on the other hand, force one to talk about the interplay between innate factors and the environment during development. A constraint is defined in terms of how the individual interacts with the environment and is unthinkable apart from such interaction. This leads researchers like Peter Marler (1991) to challenge those who “think of learning and instinct as being virtually antithetical” by speaking of an “instinct to learn.”

Many in the psychological literature take the emphasis on domain-specific constraints a step further and actually use the term “domain specific” as a surrogate for “innate.” Karen Wynn makes this point explicitly for innate knowledge: “The substantive issue concerning a given body of knowledge is ‘what is the nature of the built-in mental mechanisms that are responsible for the emergence of the knowledge?’ With regard to this question the term ‘empiricist’ typically applies to accounts positing a general learning mechanism (classically, the laws of association), while ‘nativist’ applies to accounts involving domain-specific mechanisms” (1992c, 378). Although in this remark Wynn only equates domain-specific principles with innate knowledge, in practice she also equates it with innate concepts and abilities, depending on the sort of constraint in

question. Robert Schwartz also explicitly equates innateness with domain specificity, “On the one hand, nativist theorists argue that what is brought to the task of acquisition must be *domain-specific* knowledge having particular numerical content. In contrast, nonnativists maintain that what is innate are generalized capacities and abilities, not peculiar to the domain of mathematics” (1995, 227).¹⁹ Substituting domain specificity for innateness makes sense in terms of the intuitive definition of innateness I gave above. A constraint has informational content, in that it divides actions into a class that should be performed and a class that shouldn’t. If a response to an experience is based on a domain-specific constraint, then we can think of the experience as triggering the information implicit in that constraint. Claiming that development is based on domain-specific principles thus moves the content of the knowledge into the structure of the mind and away from the structure of experience.

There are a couple more concrete arguments in favor of the equation of innateness and domain specificity. Andy Clark (1993) has argued that any rationalist (i.e., nativist) account of a body of knowledge that only specified domain-general constraints on

¹⁹ The equation of “innate” and “domain specific” is also implicit in remarks such as Frank Keil’s (1994, 235) description of theories of the development of biological knowledge: “In the empiricist view, concepts in biology, as in every other domain, emerge via general associative or inductive mechanisms.” The description of learning research in ethology given by Gallistel et al. (1991, 4) treats domain generality as the natural contrast for innateness: “Far from seeking general laws of learning, the main tenant of this creed is innately directed, or preferential learning.” Other writers do distinguish domain specificity from innateness, such as Chomsky (1980, 40 ff.) and Gelman (1993, 65).

development would collapse into empiricism. A domain-general rationalism, “would have to involve strategies which successfully go beyond the data in any domain. But how could this be? For to go beyond the data means to reach conclusions not reachable without specific pre-information. Any principles which successfully apply to any domain must therefore be exploiting information implicit in the data and/or relying on completely general facts about the structure of our universe. Mechanisms exploiting these kinds of regularity clearly fall into the empiricist camp.” Well, not really. Mechanisms relying on completely general facts about the universe only seem empiricist if you have already equated domain specificity with nativism. Indeed, in Locke’s time, many of the candidates for innate knowledge were precisely very general features of the universe, such as the idea that an object cannot both exist and not exist (1689/1964, 67).

Wynn’s argument for substituting domain specificity for innateness is that it enables one to distinguish learning from triggering more rigorously than would otherwise be possible. If we simply thought of innate knowledge as knowledge present from birth and attempted no further analysis, we might try to identify a piece of knowledge as innate by showing that an infant, not long in this world, has the piece of knowledge in question. But this strategy will never be conclusive, because it is always logically possible that the idea was learned in the short time the child had been in the world. Conversely, no matter how late an idea seems to appear, one can always say that the appropriate maturational trigger was not in place. On the other hand, Wynn claims, if “innate knowledge” is equated with “developed according to domain-specific principles,” it becomes possible to conclusively identify innate knowledge. What Wynn has latched onto here, however, is simply the

underdetermination of theory by evidence. Infant studies are no more inconclusive than any other sort of study. While a challenge can always be mounted to them, if the challenge is continuously met, the opponents are going to give up. Wynn's preferred definition of innate is in no better position.

While the notion of a domain-specific constraint is certainly very productive, there are many reasons to be dubious of simply equating it with innate knowledge. What part of a process of development governed by domain-specific principles should we take to be synonymous with innate knowledge? Are the domain-specific principles themselves the innate knowledge, or is the whole body of knowledge that develops on the basis of these principles considered innate? The latter alternative seems incautious. If we say that the whole of the body of knowledge that develops is innate, we would be essentially saying that the constraints on learning contribute the bulk of the content of that body of knowledge. But we do not know what factors besides domain-specific constraints affect development, so it would be precipitant to rule them out as contributors of content. Besides, where would we draw the line between the innate body of knowledge and the work that builds on it? Would we call the theory of groups innate because the development of counting is constrained by principles like stable ordering? It must then be the constraints themselves that are supposed to be the surrogates for the notion of innate knowledge.

The constraints themselves, however, seem more like innate concepts or abilities than they do pieces of innate knowledge. Marler's characterization of them as instincts to learn brings out the sense in which they are abilities. Domain-specific constraints are the

abilities required to engage in a learning process; they are not the thing learned. Domain-specific constraints also bear a similarity to innate concept, in that what is often required to engage in a learning process is a concept. For instance, it has been hypothesized that when a child sees an adult point to an object and say a word, the child assumes the word is a name for the object, unless there is evidence to the contrary. This constraint on the way the child acquires language is built on the possession of a concept, the concept of a class of objects. Examples like this indicate that if we are going to use “domain-specific constraint” as a surrogate for “innate capacity,” we will be better off viewing them as replacements for our notions of innate concepts and abilities than innate knowledge.

Even granted this refinement, equating domain specificity and innateness is a bad idea. Pretheoretically, the terms have different meanings, and the pretheoretic meanings of the former do not encompass all the pretheoretic meanings of the latter. In ordinary speech, we can say that the development of a cognitive ability is constrained by a domain-specific principle that was acquired some time after birth. There is nothing contradictory in saying that the stable-ordering principle is learned before a child starts to count and then governs all of the child’s later attempts to count. Really, to be an adequate substitute for innate knowledge, domain-specific constraints must be present from birth. They must be *innate* domain-specific constraints. But now all of the old problems reassert themselves. How are we to distinguish innate from learned constraints? This is especially difficult because most of our evidence for the existence of constraints is simply the observation that certain possible lines of development are not followed. We observe that children rarely violate the stable-ordering principle and conclude that it is constraining the

development of their counting. The origin of the constraint is unknown.

In general, I am not sure why I should regard a disposition to acquire knowledge in a certain way as itself being a piece of implicit knowledge, an implicit concept, or an implicit ability. My ability to acquire languages is constrained by the shape of my tongue and larynx: I will never be able to speak dolphin languages. My ability to acquire any knowledge is constrained by my short attention span and poor memory. None of these things count as implicit knowledge or implicit concepts. They are also not abilities in themselves. As aspects of my make-up which bear on my ability to learn, they do not affect how I oversee my actions, they only affect the actions I am capable of. Thus they are not aspects of my know-how.²⁰ But to distinguish constraints like a short attention span from the cognitive constraints involved in implicit knowledge concepts or abilities, we must reopen the question of what counts as knowledge in this circumstance, returning us to where we started.

The second approach I would like to criticize is much simpler than the domain-specificity approach. Why not just equate content with causation, and say that if the causes of a capacity can be found solely within the organism, the capacity is innate? The idea is quite simplistic, and no one has endorsed it flatly. After all, equating 'innate' with 'caused solely by the internal workings of the organism' means dropping the notion of triggering altogether, which leaves us with a very restricted, if not empty, class of innate

capacities. However, simple ideas have great appeal, and there is at least one definition of innateness with its roots in the causal approach. In their book *Rethinking Innateness*, Jeffery Elman and his colleagues take a stab at producing a causal definition of innateness. They call any aspect of an organism's phenotype innate if it is the product of interactions that take place solely within the organism (1996, 22). They compensate for the small number of innate capacities this definition leads to by adopting a distinction from Mark Johnson and John Morton (1991). Johnson and Morton differentiate between the *species-typical environment* and the *individual-specific environment*, and label an aspect of an organism's phenotype *primal* if it arises from interactions that take place within the organism or between the organism and its species-typical environment. Johnson and Morton's notion of primal traits seems like a fine substitute for an innate trait. Unfortunately, it comes with the same kind of rider that Leibniz put on his notion of innateness: "Note that structures arising following specific information from the species-typical environment (learning) would not be classified as being primal" (ibid., 10). Here again, our definition imports notions of triggering and learning under the guise of interactions that bear information and those that do not. Unfortunately, no technical definition of information, such as Shannon's definition in terms of entropy, will

²⁰ This is why it was important to distinguish know-how from mere regular action at the beginning of this section.

adequately capture the distinction we need to draw, because such definitions cannot distinguish between useful and useless data. Moreover, if the definition of “primal” did not exclude information-bearing interactions with the species-typical environment, it would be far too inclusive to be useful. Consider the issue that Johnson and Morton take as their topic, the development of facial recognition in infants. The presence of human faces surely must be considered a part of a baby’s species-typical environment. If we did not exclude information-bearing interactions from the definition of primal, all knowledge of faces would end up being primal. The situation is similar for the topic of this dissertation, the development of mathematical knowledge. The presence of movable, discrete objects is a part of the typical human environment. If primal knowledge encompassed all interactions with the species-typical environment, knowledge of elementary arithmetic would be trivially primal.

The Representationalist Nativist Position

The best way, I think, to develop a strong nativist position is through the notion of a representation. This is a useful route because there is already a notion of a representation in the psychological literature. C. R. Gallistel, who together with Rochel Gelman has put forward one of the prominent models of the development of mathematical knowledge, says that the brain should be said “to represent an aspect of environment when there is a functioning isomorphism between an aspect of the environment and a brain process that adapts the animal’s behavior to it” (Gallistel, 1990, 4). Gallistel intends the word “isomorphism” to be understood in its straightforward, mathematical sense. He is also

committed to the physical existence of the representing structures. He does allow, however, that the isomorphism may be mediated by a code or language (*ibid.*, 29). Although there is an isomorphism between the real numbers and the weights of physical objects, the symbol '5' is not heavier than the symbol '4'. The isomorphism is only present when the language of number is taken into account. Similarly, the brain structures isomorphic with the environment may only be isomorphic when interpreted as symbols in a language. Admittedly, there is much that is vague in this account of representation. For instance, it is not clear whether the isomorphism in question involves types or tokens. Gallistel does emphasize the systemic nature of isomorphisms: "the statement that the number 10 is isomorphic to some particular mass is meaningless. It is the system of number that is isomorphic to the system of masses" (*ibid.*, 24). However this is not equivalent to saying that the isomorphism must be between types. A token with five parts can be isomorphic to a collection of five items—the elements can be placed in one-to-one correspondence and the same operations and relations can be defined for each. Seeing this is no harder than seeing that a type with five parts can be isomorphic to a type of collection. Either tokens or types can exhibit the requisite systemicity. This ambiguity in Gallistel's definition will come up crucially in Chapter 4, where it plays a key role in the refutation of representational nativism.

It is a natural extension of Gallistel's notion of representation to say that knowledge of some aspect of the world is innate if one is born with a representation of that aspect of

the world. Conversely, if such a representation does not appear until after some interaction with the world, the body of knowledge can be said to be learned.²¹ For skill knowledge, this definition would be modified to say that relevant isomorphism would be between a way of working with the world and a brain structure, with the further caveat that the isomorphism must be relevant to producing the skillful action. Innate concepts could be defined using isomorphism between brain structures and the properties that unite the objects falling under the extension of the concept. This representational way of defining innate capacities can be seen as a straightforward way of adding rigor to the idea that a capacity is innate if its *content* could be viewed as lying on the side of the brain, rather than on the side of experience. The notion of content has been made more definite through the idea of a representation.

It is worth noting that this definition is fairly limiting: many routes of development that psychologists are inclined to call the flowering of innate knowledge are not covered by the representationalist umbrella. This problem is brought out nicely by the typology of innate structures proposed by Elman et al. (1996, 24 ff.). According to Elman et al. three features of one's make-up lead to knowledge being innate. First, the brain might possess innate representations of the sort we have been discussing. Second, the general architecture of the brain might guarantee that certain knowledge will develop by fixing

²¹ It is a further natural extension of these remarks to say that if a representation never exists in the brain, then there is no knowledge at all. But we do not need to open that can of worms here.

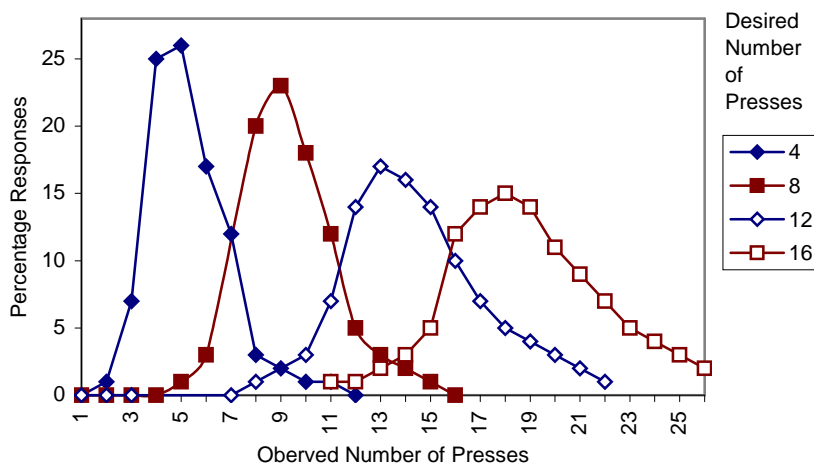
the kind of information it receives and the solutions it can consider for the problems posed by that information. For instance, Elman and his colleagues point out that the sort of representation formed by a region of the cortex is dependent on the sort of information it receives. The visual cortex takes the shape it does because it receives information from the retina. Further specification of the sort of input a cortical region receives could fix the range of possible representations to the extent that we would say, for instance, that our sight is governed by an innate object concept. Third, a timing mechanism might guarantee that the brain acquires certain knowledge. For instance, if an area of the brain only responds to input during a specific stage of development, the structures that form in that region will be shaped by the information available during that period of time. Mechanisms like this could underlie critical periods for parental imprinting to language learning. If we adopt a representational definition of innateness, knowledge generated by timing or architectural mechanisms will not be considered proper innate knowledge.

Why should we believe that there are innate representations of this sort? Over the last several decades, developmental psychologists have been accumulating a great deal of evidence that they think shows that some mathematical capacities are innate, and have put forward several models of that data, one of which involves an innate representation as we have just defined it. In the remainder of this section, we will look at that data and the model that involves an innate representation.

The data for mathematical innateness can be roughly classified under four headings: animal studies, infant studies, childhood studies, and adult studies (including studies of response times and error rates, and studies of selective brain damage, and EEG, PET, and

MRI scans of adults performing mathematical tasks). Animal studies include work with creatures with no linguistic ability, such as rats and pigeons, as well as animals that can be trained in human speech, like chimps or parrots.²² The most important studies with rats follow a design developed by Francis Mechner (1958). In these experiments, rats are trained to press a lever a certain number of times, N , before pressing a second lever to get a reward. After a learning period, the rats were able to approximate the correct number of lever pulls, but their approximations got steadily worse as N increased (see Fig. 1). Later refinements of the experiment ruled out the possibility that the creatures were responding to duration, rather than number (Mechner and Guevrekian 1962; Laties 1972; Fernandes and Church 1982; Meck and Church 1983). These experiments also indicated that rats

Fig 1: Counting Ability in Rats. Redrawn from Mechner (1958)



could count stimuli as well as responses (Fernandes and Church 1982), can generalize across sensory modalities (ibid.; Meck and Church 1983), and

²² For an overview of experiments in animal understanding of number, see Gallistel (1990, ch. 10).

can even respond to the sum of stimuli given in two different modalities (Meck and Church 1983). Although animal counting was carefully teased apart from animal perception of duration, several important parallels remained between the two faculties. The probability curve for a right response given different sizes of collections matched the probability curve of a right response given durations of different length (ibid.). Both curves are also effected by methamphetamine in a parallel fashion (ibid.). Experiments with animals with some linguistic ability are based around teaching the creatures to use number words in the course of their linguistic training. A Japanese chimp named Ai, for instance, apparently acquired the numbers one through six in the course of learning an artificial language of symbols on a keyboard (Matsuzawa 1985). Similar abilities have been found in a creature one might be less inclined to attribute real linguistic ability to, the African gray parrot (Pepperberg 1987). In all these experiments, whether with rats or chimps, animal number sense was approximate, error prone, and limited to small numbers. Specifically, animal number sense displays a *magnitude effect*—the error rate increased linearly with the size of the numbers involved—and a *distance effect*—the closer two numbers were to each other, the harder they were to distinguish.

Animal studies are really only indirectly relevant to the study of mathematical knowledge in humans. They enter into model building via the assumption that the human brain was built on top of older, more primitive capacities. More direct evidence for the innateness of mathematics comes from the study of human development. The most

striking studies in human development come from a pair of protocols developed in the 1960s for judging the cognitive abilities of infants.²³ These studies are based on paying close attention to the way infants attend to different aspects of their environment. The first of the classes of studies are the *habituation studies*, which consist of two steps. First, an infant is exposed to an image until her eyes cease to focus on it, wandering to other sights. Then the baby is exposed to an image that varies slightly from the first. If her attention returns to the new image, the researcher concludes that the infant can distinguish between the two images. If not, then the images are inferred to be indistinguishable. A related class of studies are the *preferential looking studies*. Preferential looking studies forgo the initial, habituation phase of habituation studies. Instead, two images are presented simultaneously. If infants consistently look longer at one of the images, it is again concluded that they can distinguish them. Habituation studies have been used to show that newborns (Antell and Keating 1983) five month olds (Starkey and Cooper 1980), and 10 month olds (Strauss and Curtis 1981) are capable of discriminating between collections as large as three and sometimes four. All sets of four or larger, however, are apparently indistinguishable (Starkey and Cooper 1980; Strauss and Curtis 1981). Related studies have shown that infants can discriminate between numerosities across sensory modalities. Starkey et al. (1983, 1990) played recordings of two or three

²³ There is no single place for a good overview of studies on infant mathematics, although Wynn (1992b) makes a start.

knocks to six to nine month old infants and then showed them a picture of two objects and a picture of three objects simultaneously. The babies showed a preference for looking at the picture that matched the number of knocks they heard.²⁴ Finally, Karen Wynn (1992a) believes she has found evidence that infants have expectations regarding the sums and differences of the numerosities they perceive. A typical experiment she has performed began with a series of infants observing two objects being placed one after another behind an opaque screen. In half of the trials one of the objects was then surreptitiously removed. When the screen was removed, the infants who only saw one item showed markedly longer looking times than the infants who saw both the original items. Here the limit of infant's ability was even lower than in the number discrimination tests: Wynn's work dealt with no collections larger than two. Interestingly, infants seem to identify objects in these collections simply by their trajectory and not their size, color, or shape (Xu and Carey 1996). If two very different objects, such as a red ball and a yellow ducky, are hidden behind a screen, and then slid out from behind the screen into the infants vision one at a time, the infant will show no surprise when the screen is dropped to reveal a single object. She will not change her attention pattern in any way.

²⁴ There is contradictory evidence here. David Moore and his colleagues (1987), trying to replicate Starkey et al.'s experiment, obtained the opposite results. They found that infants showed a preference for collections that do not match the knocks they hear. Gelman (1990) gives an odd reply to this study, saying that it doesn't matter which set is preferred. As long as there is a preference we can infer that infants can distinguish the two sets. However, taken together, the studies show that there is not a consistent preference. To my mind the question of cross-modal sensitivity to number is still open.

The infant will change her attention pattern, however, if she is allowed to see both of the toys move from behind the screen at the same time before the screen is dropped, indicating that the child can use spatiotemporal data to individuate objects.

Infant studies put a lot of stock in very small differences in human behavior. Studies of older children learning counting and basic arithmetic offer less subtle forms of evidence about the development of mathematical knowledge. Interestingly, the very first studies of children performed were taken to show that mathematical knowledge is not innate. Childhood studies were pioneered by Jean Piaget, the father of developmental psychology.²⁵ Piaget's most famous experiments with children involved number conservation. A typical experiment in number conservation would present a child with two rows of objects lined up so that it was clear that there were the same number of objects in each row. The experimenter would ask the child if she knew that the rows held the same number of objects. After the child confirmed that they did, the experimenter would spread the items in one of the rows farther apart. Children under the age of six or seven would inevitably report that the longer row now had more objects in it. Older children, on the other hand, would recognize that the change had not increased the number of objects, and would often find the experimenter's questions rather silly. Results like these lead Piaget and his colleagues to conclude that children did not grasp very important aspects of the concept of number before the age of six or seven.

²⁵ For an overview of Piaget's work see Piaget (1952).

Piaget's findings regarding conservation have remained essentially unchallenged. Evidence has been found that children begin conserving slightly earlier than Piaget thought, however as Rochel Gelman and C. R. Gallistel note, "The failure of children younger than five to conserve...is one of the most reliable experimental findings in the entire literature of cognitive development" (1978, 1). More recent studies of childhood mathematical knowledge have not refuted Piaget, so much as they have gone beyond him. They indicate that while children may not conserve, they do possess a great deal of other knowledge relevant to mathematics. These studies have focused on children as they learn to count. Counting, in the developmental literature, is typically assumed to be defined by five principles, which were originally set out by Gelman and Gallistel in their book *The Child's Understanding of Number* (1978). The first three of Gelman and Gallistel's principles are called the how-to-count principles, because they define the process of counting itself. They are the principle of *one-to-one correspondence*, which defines how a set of objects is to be coordinated with a set of words or gestures; the *stable-ordering* principle, which says that whatever number tags are used, they must be used in the same order each time; and the *cardinality* principle, which says that the last tag applied to a set represents the numerosity of the whole. The remaining two principles are metarules, called by Gelman and Gallistel the what-to-count principles, because they define how the first three principles are to be applied. They are the principle of *item indifference*, which

says that any array of objects can be counted, and the principle of *order irrelevance*, which says that the items of a collection can be counted in any order.²⁶

In *The Child's Understanding of Number* and subsequent publications, Gelman and Gallistel have argued that the knowledge of the counting principles precedes knowledge of how to count and is what allows for that mastery to develop. They certainly present a great deal of evidence that children understand these principles at a very young age, before they are able to count consistently and accurately. With regard to the one-to-one principle, for instance, they note that children who have not yet mastered counting, “seldom produce totally asynchronous counts; in general they point to and touch a single item...and state a single numerlog” (1978, 205). Trouble synchronizing number words and objects sometimes occurs at the beginning or end of counts—the child might start reciting words too early or continue to recite them after she is done pointing to objects—but during the bulk of the count, the process of pairing words and objects goes smoothly. Evidence for early knowledge of the stable-ordering principle is also quite strong. Children who have not yet memorized the sequence of number words will generally use a sequence of their own devising, e.g., “one, two, six,” which they apply consistently (ibid., 90). Idiosyncratic lists generally consist of standard number words, and if other tags are used, they are quite often words from other invariant lists, such as letters of the alphabet

²⁶ I recognize that as a philosophical analysis of counting, these principles are lacking. I discuss them because they have become standard in the developmental literature.

(*ibid.*). The abstraction and order irrelevance principles are also evident in young children: “Over and over again, we found that even our very youngest subjects were willing to classify together and count diverse items”; “In several of our studies we have asked children to count the number of items in a given set several times. Usually we rearranged the items...in our analysis of counts across rearrangements, we seldom found any evidence of children attempting to keep giving the same numerlog to a particular item. The finding was true no matter what the age of the child” (*ibid.*, 138 and 217). The principle for which evidence was lacking was the cardinality principle. Children often repeat the last number of a count, or say it with particular emphasis, as if they attached special significance to it. However, this can be chalked up to factors like mimicking of adult behavior (*ibid.*, 207 ff.).

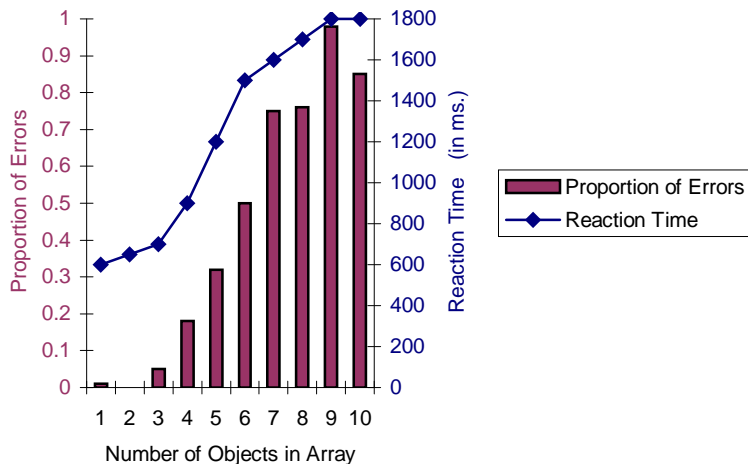
Children thus seem to be aware of the counting principles, even before they are able to competently employ them in counting. But what role do these principles play in facilitating the acquisition of counting? According to Gelman and Gallistel, the counting principles constitute a domain, which they define technically as a set of entities combined with operations on those entities (*ibid.* ch. 208 ff., for a more precise statement see Gelman 1990, 81). The counting principles form a domain because they highlight certain sensory inputs, marking out the pool of data used to form representations of numerosity, and then define a set of operations on those representations, the operations used in counting. Evidence that the counting principles highlight parts of the sensory input can be found in the patterns of attention observed in the infant habituation studies. The methods children use to separate number words from object names when they are learning to speak

also indicate that the counting principles mark out a unique set of sensory inputs (Gelman 1990). When an adult points to an object and says a word, the child assumes, among other things, that the word is a name for the object (Markman 1990). When an adult points to several objects of the same kind in sequence, reciting a new word each time, the child recognizes this as a different sort of word game, and is prepared to assimilate the adult's words into the class of number words. Evidence that the counting principles meet the second part of Gelman and Gallistel's definition of a domain, that they are actually used to define operations on the set of inputs they highlight, can be found in the way children correct themselves when they are learning to count (Gelman and Gallistel 1978, 93). Even very young children will, without any prompting, correct their attempts to count so that that they conform with the counting principles.

The final class of evidence the mathematical nativist is concerned with consists of studies of adult competence in mathematics. These studies include experiments with the response and error rates of adults faced with basic arithmetic problems, cases of selective loss of mathematical capacities in adults with brain lesions, split-brain patients, and EEG, PET, and MRI scans of adults performing arithmetic tasks. Two interesting phenomena can be observed in the response time and error rate studies. The first deals with our ability to perceive the size of small collections. Introspectively, it often seems like we perceive the numerosity of small collections *immediately*, whereas larger collections require a bit of mental counting. Scores of reaction time and error rate studies, dating back to Jevons

(1871), show that there is something to our introspective feelings here.²⁷ Since Kaufman et al. (1949), this process of quickly grasping the size of collections has been called *subitizing*, from the Latin *subitus*, sudden. Various theorists put the upper limit on subitizing anywhere between three and seven. The most frequently cited study from contemporary times is Mandler and Shebo (1982), who used a tachistoscope to flash

Fig. 2: Reaction Time and Error Rate in Judging the Size of Collections. Redrawn from Mandler and Shebo (1982)



collections of various sizes before subjects for a set period of time. The result are given in Figure 2. They seem to show two discontinuities: the slope for the reaction time is shallow between one and three dots, nearly linear between

three and six, increases by about 200 ms per dot, and levels off after six dots. Errors start appearing with three dots and become overwhelming around seven dots. This often leads researchers to think that a separate process of subitizing governs the perception of

²⁷ For an overview of this literature, see Dehaene (1992) or the opening discussions in Trick and Pylyshyn (1993), Gelman and Gallistel (1991), or Mandler and Shebo (1982).

collections of three or less, and that a process of unconscious counting governs the estimation of larger collections, but that this counting process loses its effectiveness around six or seven. On the other hand, the presence of a slope at all for collections below three militates against this interpretation. If there were really a holistic process governing the perception of collections this small, the response times should be perfectly flat.

The other phenomenon that crops up in error rate and reaction time studies involves the distance and magnitude effects observed in adults comparing, adding, or multiplying numbers.²⁸ One would not expect these effects to crop up if people were calculating using only the linguistic or logical properties of numbers, but they show up persistently and in a variety of tests. When adults are asked to determine the larger of two numbers presented as Arabic numerals, their reaction times and error rates show a magnitude effect: larger numbers require more time and produce more errors. The digits '8' and '9' are taken longer to compare than the digits '2' and '3'. Such comparisons also obey a distance effect. It takes longer to compare '8' and '9' than it does '1' and '9' (Moyer and Landauer 1967). This result continues to hold for two digit numbers (Hinrichs et al. 1981 ; Dehaene et al. 1990). It takes longer to compare '69' and '71' than it does to compare '69' and '78', even though in both cases one need only look at the tens column to make the comparison. Strategies based on just looking at a few digits only kick in with numbers of

²⁸ For an overview of the literature on the magnitude and distance effects in number comparison see Dehaene et al. (1990) or Dehaene (1992). For an overview of the

three or more digits (Hinrichs et al. 1982; Poltrock and Schwartz 1984). Interestingly, the exact function that governs the reaction times in number comparison is a logarithmic equation,

$$(1) \text{ RT} = a + k \log [L/(L - S)],$$

where RT is the reaction time, L is the larger number, S is the smaller number, and a and k are constants. This same equation was originally shown by Welford (1960) to explain the reaction times in comparisons of physical magnitudes, such as distance or musical pitch. The magnitude effect also shows up in the use of addition or multiplication, even though most people are almost certainly answering these questions using rote memory. The effect obeys the same function in both multiplication and addition, a function that was initially thought to be linear (Parkman 1972) but was later shown to be exponential (Ashcraft and Battaglia 1978). A notable exception to this rule is that reaction times for problems where both operands are the same (e.g., 9×9) are much lower (Ashcraft 1992, 80). Finally, it is worth noting errors in multiplication and addition usually take the form of the answer to a different but related problem, such as the answer '32' for the question 'what is 4×6 ', which is what you would expect from a rote memory task (Campbell and Graham 1985).

magnitude effect and related phenomena in arithmetic operations like addition, see Ashcraft (1992).

Studies of brain-lesioned patients begin with the observation that such patients often display very specific deficits, such as the inability to read verbs or to acknowledge anything having to do with the right side of their body. This leads researchers to ask, “What must normal cognitive architecture be like such that damage to it can produce this pattern of disability?” Researchers in this area have generally been careful not to make the kind of rash generalization often attributed to them, namely, the claim that if a function can be selectively disabled, it must be an autonomous part of the normally functioning brain.²⁹ Nevertheless, lesion studies must rely on at least two major assumptions. First, they must assume that the brain does not massively reorganize itself after physical trauma, making it impossible to infer anything about normal function from posttraumatic functioning. This assumption is crucial, because there is ample reason to think that the brain can reorganize itself in just that fashion. Second, because of the small number of patients displaying any given deficit, researchers must assume that the cognitive architecture is uniform for all people with regard to the abilities in question.

If one accepts these assumptions, there is a wealth of interesting data about mathematical cognition from brain-damaged patients with *dyscalculia*—disability in the use of number.³⁰ Dyscalculias can be divided into two classes. First there are those

²⁹ For an overview of the debate over lesion studies see Caramazza and McCloskey (1988).

³⁰ For an overview of the literature on dyscalculia see McCloskey (1992) or Levin and Spiers (1985).

dealing with the production and comprehension of number—essentially input and output errors. Second there are those dealing with the internal processing of numbers.³¹ Deficits of the first kind cover the production and comprehension of spoken numerals, written Arabic numerals, and written number names, as well as the ability to translate between these three codes. One of the oldest observations in the literature on dyscalculia deals with the ability to translate between the Arabic and spoken code. There are many reports of brain-damaged patients who can no longer read regular words, single letters, or written number names aloud, but retain the ability to read Arabic numbers (e.g., Déjerine 1892; Greenblatt 1973; Anderson et al. 1990). The reverse impairment has also been found, although it is more rare (Cipolotti 1995; Cipolotti et al. 1995). The case described in Lisa Cipolotti et al. (1995) is interesting because the patient was only unable to produce a verbal number when presented with a Arabic number—she had no difficulty verbally answering verbal arithmetic questions, for instance. Similar observations can be made about the ability to write Arabic numerals to dictation (Anderson et al. 1990; Deloche and Seron 1982a, 1982b, 1987). A more interesting observation is that production impairment can be limited either to the syntax of numbers or to the ability to recall specific number words (Singer and Low 1933; McCloskey and Caramazza 1987; Cipolotti et al. 1994). A

³¹ Many writers make a different distinction, first proposed by Berger (1926). This line of thought distinguishes *secondary dyscalculia*, which arises from impairment of a general purpose faculty like short-term memory, from *primary dyscalculia*, which is a disruption of a specifically mathematical faculty. I will not adopt that distinction here, so as not to prejudice against the possibility that there are no specifically mathematical faculties.

patient with impaired syntax would write '1000,945' when she hears 'one thousand nine hundred and forty five', whereas a patient with a lexical impairment might write '2,962'. Deficits in comprehension also include deficits in ability to judge the numerosity of collections of objects. Here the significant results are cases where the ability to count objects can allegedly be dissociated from the ability to subitize (Dehaene and Cohen 1994; Sokol et al. 1994).

Deficits in number processing generally revolve around the ability to perform arithmetic operations. Here the most common observations involve the loss of the ability to retrieve basic arithmetic facts, like $5 \times 6 = 30$, with the preservation of some ability to reason to the correct answer (Warrington 1982; Sokol et al. 1989; Hittmair-Delazer et al. 1994; Dehaene and Cohen 1997). The patient described by Hittmair-Delazer et al. 1994, for instance, would multiply by five by multiplying by 10, and then dividing by two. In another paper (Hittmair-Delazer et al. 1995), Hittmair-Delazer and colleagues describe a case in which the patient is actually able to state the rules he is using in algebraic form. Other patients who have difficulty recalling arithmetic facts can still use the algorithms for multidigit addition and multiplication correctly, although they plug the wrong values into those algorithms (Sokol et al. 1991, 1994). The opposite dissociation has also been noted—some patients are able to recall arithmetic facts, but show no ability to reason with them (Cohen and Dehaene 1996; Dehaene and Cohen 1997; Sokol et al. 1994). It is also possible to lose the ability to recall certain arithmetic facts but not others. For instance, Scott Sokol et al. (1989) report on a patient whose dyscalculia focuses on solving problems involving 7 or 8. The most important result in the literature on deficits

in number processing is the presence of an approximate number sense in patients who have lost their ability to perform exact calculations. Patients will be unable to say if $2 + 2$ is 4 or 5, but they will be quite sure it is not 13 (Warrington 1982; Dehaene and Cohen 1991).

Some researchers will leave information about the location of a lesion out of their report because they believe that cognitive science should only be concerned with the design of the mind, and not its implementation in the brain. However, there are interesting data to be gleaned from the location of various lesions. The woman described in Anderson et al. (1990) whose alexia and agraphia did not extend to Arabic numbers had a circumscribed lesion in the left premotor cortex, Brodman's area six, right behind Broca's area. The few cases of the converse dissociation, noted in Cipolotti (1995) and Cipolotti et al. (1995), had damage to the left parietal cortex. Difficulties in recalling arithmetic facts are associated with damage to the basal ganglia, which are generally implicated in the production of rote behavior. Patients with damage isolated to this area generally retain the ability to produce and comprehend number and perform other arithmetic tasks. (The first patient in Dehaene and Cohen 1997 as well as the patient in Hittmair-Delazer et al. 1994 fits this pattern.) On the other hand, damage to the left inferior parietal region, especially around the angular gyrus or Brodman's area 39, results in a deeper acalculia, including loss of arithmetic reasoning and calculation. (The second patient in Dehaene and Cohen 1997, as well as the patient in Cohen and Dehaene 1996, fits this pattern.) Split-brain patients, individuals who have had their corpus callosum severed to control epilepsy, are also useful in localizing mathematical ability. Split-brain

studies indicate that either hemisphere is capable of deciding whether two numbers are the same or different and deciding which is larger (Seymour et al. 1994). However, only the left hemisphere is able to use verbal numbers or, more significantly, answer simple arithmetic problems in any format (Gazzaniga and Smylie 1984). The final kind of data useful for localizing arithmetic functioning come from brain scan studies, however not all the researchers performing this kind of study have asked the kind of questions we are interested in. There are, for instance, several studies of patients engaged in tasks like counting backwards by threes (Roland and Friberg 1985; Appolonio et al. 1994). These studies showed extensive activation of the prefrontal cortex. However, the task studied has long been acknowledged as more a test of short-term memory, attention, and planning than it is of any component of arithmetic ability. Stanislas Dehaene and colleagues (1996) have conducted a more interesting study. Subjects were shown two Arabic digits and asked to either decide which is larger or multiply them in their heads. PET scanning for both tasks showed a number of areas of activation believed to be due to either reading the stimulus or speaking the answer. Strangely, the comparison task did not reveal any activation apart from these areas. There was some activation of the left and right inferior parietal cortex, but not at significant levels. There were significant levels of activation in these areas for the multiplication task, as well as activation of the left fusiform/lingual region, as well as the left lenticular nucleus.

Most readers by now are asking what all this has to do with the philosophy of mathematics. The reason all these data are important is that they can be used to argue that there is a mechanism that represents mathematical facts, which we share with mammals

like rats and chimps, which is operating in the brains of infants, which guides children in learning to count, which underlies much of adult competence in mathematics (perhaps everything short of arithmetic tables learned by rote), and which requires the functioning of something in the left inferior parietal region, especially around the angular gyrus or Brodman's area 39. There are a couple versions of this hypothesis, and different versions work better with different definitions of innateness. Right now I will look at the version that came out of animal psychology and consider it in light of the representationalist definition of 'innate'. When we turn to the weak version of innateness, I will consider an alternative, connectionist model.

The animal model of counting began its life as an attempt by John Gibbon and Russell Church to model the perception of duration in rats. (For an overview of the proposal, see Gibbon and Church 1984.) The key component of this model was an internal clock consisting of a pacemaker, which regularly generates impulses; an accumulator, where these impulses are stored; and a switcher, which controls whether the impulses from the pacemaker pass into the accumulator. Images of the state of the accumulator can then be taken up into working or long-term memory and used in decision making. By fiddling with the values assigned to the rate of the pacemaker and the time it takes for the switch to open and close and making some assumptions about animal memory and decision making, one can approximate the data received from experiments with rats' perception of duration. Warren Meck and Russell Church (1983, also Church and Meck 1984) noticed that this timing mechanism can account for discrimination of number as well as duration, if the switch is allowed to function in a number of different

“modes.” In *run mode* the switch remains open until a different impulse closes it, in *stop mode* the switch remains open only as long as a given stimulus is present, and in *event mode* the switch stays open for a fixed period of time when an event occurs. The first two modes measure duration in different ways, but the third is a measure of number. By passing a fixed number of impulses into the accumulator every time an event occurs, the mechanism is effectively keeping a tally of the number of times that event has occurred. If the same mechanism is used to judge both duration and number, it would explain why the two capacities are distinguishable, yet function in a parallel fashion. As with the duration model, it is possible to generate accurate quantitative predictions by assigning the right values to various parameters of the model (Meck and Church 1983). In this case, a key assumption is that the switch of the internal clock stays open for 200 ms for each item counted.

Meck and Church’s animal counting mechanism was picked up by Wynn (1992b) as an explanation for infant arithmetic ability. Infants recognize that a stimulus contains a novel number of dots because they can compare the current state of an internal accumulator to the record of a past state. The model was especially good, Wynn noted, at explaining her experiments with infant arithmetic. Because the representation of number in the animal counting mechanism is itself a magnitude, it is easy for it to be combined with other representations to form the sum of two numbers in internal arithmetic. A more abstract code would require additional rules that say that representations of a certain kind can be combined in ways that essentially follow the Peano axioms. Wynn (1992b) also noted that the animal counting mechanism can assist children in learning verbal and

written counting. If such a mechanism were present from birth, learning verbal numbers would simply be a matter of mapping verbal and written numbers onto the states of the internal accumulator. Gelman and Gallistel (1992) concur, although on their account, the mapping is not onto states of the accumulator per se, but onto the analog magnitudes that are built up in the accumulator and stored in working or long-term memory. Gelman and Gallistel also note that the preverbal counting mechanism obeys their how-to-count principles. It obeys the one-to-one principle because a single burst of impulses is released for a single object; it obeys the stable order principle because the states of the accumulator form a fixed sequence of magnitudes; and it obeys the cardinality principle because the final state of the accumulator is used to represent the numerosity of the set counted. Therefore, if the preverbal counting mechanism mediates children's acquisition of counting, it would explain why children never violate the how-to-count principles.

Gelman and Gallistel go further still and attribute a role in adult computation to Meck and Church's preverbal counting mechanism. In their essay "Subitizing: The Preverbal Counting Process" (1991) they argue that subitizing is simply the preverbal counting process at work. This would explain, they argue, the increase in reaction times for identifying collections of three objects or fewer—the range of collections that we are supposed to be able to subitize. Gelman and Gallistel's "Preverbal and Verbal Counting and Computation" (1992) argues further that analog magnitudes of the sort generated by the preverbal counting mechanism play a role in adult comparison of two numbers. When a person is presented with a pair of numbers, the quantities are mapped onto internal magnitudes, which are then compared. These same magnitudes, they claim, are used to

represent physical magnitudes, such as length, which is why reaction times for both magnitude and number comparison are governed by the same function. The map between verbal numbers and internal magnitudes is linear, but the exact size of these magnitudes is variable, and the variation increases as a scalar function of their size. Reaction times for both magnitude and number comparison are supposed to be governed by Welford's logarithmic equation (1) because of the difficulty in dealing with more variable magnitudes. Finally, Gelman and Gallistel argue that use of preverbal magnitudes mediates knowledge of the addition and multiplication tables. One solves a simple addition problem by mapping the addends onto preverbal magnitudes, combining those magnitudes along a mental number line, and mapping the result back to a verbal or Arabic number. Solving simple multiplication problems works similarly, except that the preverbal magnitudes are placed along the edges of a two-dimensional mental table with the answers to the multiplication problems represented as squares in the table. Again the variability of the preverbal magnitudes is supposed to explain the magnitude effect in the reaction times.

The model also explains why the same function governs the magnitude effect in both addition and multiplication. Both operations require the same processes of mapping, combining, and mapping again. Gelman and Gallistel can also explain the reduced response time for problems where both operands are the same: these problems require one less mapping from verbal or Arabic numbers to preverbal magnitudes. The pattern of errors in multiplication problems can be explained by the arrangement of the mental multiplication table. Mistaken answers are often answers to related multiplication

problems because these answers lie close to the correct answer on the mental multiplication table. Finally, Gelman and Gallistel also believe that their model is “at least compatible” with the data on specific arithmetic deficits. A weakening of “the association between a magnitude on the number line or a position in the number field” could explain why arithmetic facts can be selectively impaired (*ibid.*). Also, increased variability in preverbal magnitudes can explain cases where precise knowledge of arithmetic facts has been lost but approximate knowledge is retained.

The existence of the preverbal counting mechanism can be used to frame an argument for the innateness of mathematical capacities. Wynn (1992b) makes something like this argument, directing it explicitly against the empiricist theories of mathematics developed by Kitcher and Mill. Unfortunately, she fails to make the argument as convincing as it can be. Rather than arguing for the existence of the preverbal counting mechanism and then using the preverbal counting mechanism to argue for the innateness of mathematical capacities, Wynn claims that the empirical research she and others have done directly shows mathematical capacities to be innate. The existence of the preverbal counting mechanism is an additional hypothesis that is not necessary to refute Kitcher and Mill. Wynn does not distinguish between the mathematical capacities I have distinguished—concepts, abilities, and knowledge—but her claims can be seen as applying to all three. Basically, she reviews the experiments on animal and infant arithmetic and says, “The experiments reviewed above show that human infants and other animals possess a sensitivity to numerosity, and an ability to determine the results of simple arithmetical operations. The fact that these abilities are evident at a very early age in human infancy

suggests that we are innately equipped with such knowledge, rather than learning it through induction over experience” (ibid., 329). When Wynn claims that infants and animals are innately sensitive to numerosity, she essentially is claiming that they possess the concept of number. The claim that infants are innately able to perform addition seems to be a claim about the existence of innate knowledge and innate abilities.

Wynn’s attempt to deduce the existence of an innate number concept and innate arithmetic knowledge and ability directly from the empirical evidence runs into problems because she defines innateness as domain specificity. The empirical evidence alone simply does not point to the existence of mechanisms specific to mathematics. There are alternative explanations based on principles from other domains, in particular, the ability to use abstract concepts of objects. The main studies Wynn relies on are the preferential looking studies that show infant sensitivity to numerosity and her own habituation studies indicating infant arithmetic ability. David Galloway (1992) has argued that the sensitivity to numerosity manifested in the habituation studies is really a part of the concept of the objects involved. Part of being able to identify an object is to be able to distinguish it from objects of the same type. If I can tell that a doll is a doll, then I can already separate it from the other dolls in the experiment. Therefore, “if we attribute to the animal or child an ability to think of dolls, or beachballs, or whatever it is that is used in the experimental setup, we are *already* attributing an ability to tell one doll from two, two from three, and so on for at least small-sized aggregates of dolls” (ibid., 350). Robert Schwartz (1995) makes a similar point, writing, “The numerical sensitivity infants possess may be part and parcel of a more wide-ranging cognitive capacity to organize the world in terms of

general concepts, concepts that require mastery of principles of individuation or divided reference” (ibid., 235). Aspects of the object concept can also explain the habituation studies conducted by Wynn that purport to show infant sensitivity to arithmetic facts. Wynn’s work, you will recall, involved hiding two objects behind a screen one at a time, and then surreptitiously removing one of the objects. Increased looking time in this seemingly impossible situation was supposed to indicate that children are sensitive to the fact that $1 + 1 = 2$. But it could as easily be explained by the ability to keep track of multiple objects when they are out of sight. Thus the sensitivity observed need not be the product of a domain-specific mechanism.

Wynn has responded to Galloway by denying that one needs to be able to distinguish between individuals of the same type in order to have a concept of that type. It is empirically possible, she points out, that infants could distinguish dolls from beachballs without being able to individuate the mass of doll in front of them into individual dolls. This misses the point, however. These features may be separable, and there may even be good reason to say that they are not both necessary for the concept of an object. However, this does not mean that when one is able to individuate the mass of doll before one’s eyes that this capacity is anything besides a part of one’s general ability to deal with abstract concepts. Wynn’s second response is to point out that habituation studies have shown that infants have a sensitivity to numerosity across different kinds of items and sensory modalities, indicating that they can abstract numerosity from the notion of the objects involved. However, this would also occur if the innate principle at work here were part of one’s general ability to use any abstract concept. The ability to recognize that the number

of entities has changed across sensory modalities could be the result of sensing that one's ability to use general concepts is being used in a different fashion. One might be tempted to say at this point "but that's just what number *is*," but saying so would be to concede that number is not governed by domain-specific principles. It is governed by the principles of objecthood. For Wynn, this would concede that game.

A better way to formulate the argument for innateness would be to argue for the existence of the preverbal counting mechanism as the best explanation for the empirical evidence and then claim that the preverbal counting mechanism is a form of innate mathematical capacity. We have already completed the first part of this argument by reviewing the evidence that led Gelman, Gallistel, and others to claim that the preverbal counting mechanism explains animal numerical abilities, infant sensitivity to number, the patterns of childhood acquisition of counting, and the reaction times and error rates of adults performing arithmetic tasks. All that remains is to argue that the preverbal counting mechanism counts as an innate mathematical capacity. I propose to do that using the representational definition of innateness. Recall that we followed C. R. Gallistel and said that the brain represents a feature of the world if there is an isomorphism between that feature of the world and some functionally relevant portion of the brain. This isomorphism may be mediated by language, but the brain mechanism itself must be physically instantiated. A representation is innate simply if it is present at birth. Innate knowledge is an isomorphism between a brain structure and a fact, an innate ability is an isomorphism between a brain structure and a way of acting in the world, and an innate concept is an isomorphism between a brain structure and the properties that unite the

objects falling under the extension of the concept. It is clear that the preverbal counting mechanism is a representation of a mathematical ability, namely, the ability to count. If we accept Gelman and Gallistel's counting principles as a rough definition of counting, then they have shown that the preverbal counting mechanism is isomorphic with the process of counting by showing that it obeys their counting principles. It is also fairly easy to see that the preverbal counting mechanism is a representation of a mathematical concept, the concept of number. In fact, the magnitudes that pass from the pacemaker to the accumulator are explicitly labeled *numerons*, or representations of numerosity. I would assume that the properties uniting the objects falling under the concept "numerosity"³² are given by the Peano axioms. The magnitudes of the preverbal counting mechanism are manipulated in a way that obeys these axioms. So it is reasonable to say that in the "language" of the preverbal counting mechanism, the numerons are isomorphic to the properties that unite the objects falling under the concept "numerosity." The case of mathematical knowledge is a little rougher. While the mechanism may be isomorphic to the process of counting, individual facts of arithmetic are not immediately represented. Instead, a representation of them is formed when one deals with objects in the world. For instance, if one sees two groups of two objects placed next to each other in the world, two groups of two impulses will pass into a mental accumulator, giving the result of four. On

³² I am not making any assumptions here about what these objects are. They could be physical collections or Platonic objects.

one level, this can be seen as a process of learning. After all, a representation is formed in light of an experience. Saying this, however, would lead to the strange consequence that an arithmetic fact is relearned every time a human or animal draws on the preverbal counting mechanism. A simple alternative is to say that the preverbal counting mechanism represents arithmetic facts because it has the *potential* to produce representations of any arithmetic fact. This claim is rather vague, but it will have to do for now. As we shall see in Chapter 4, the difficulty in identifying exactly what fact the preverbal counting mechanism represents is the downfall of the representationalist nativist position.

The Intuitionist Apriorist Position

The term ‘a priori’ gets applied to a slightly different range of entities than the term ‘innate’. Like the term ‘innate’, ‘a priori’ has been applied since the Enlightenment primarily to knowledge and concepts. ‘A priori’ gets its current definition almost entirely from Immanuel Kant,³³ who described a priori knowledge as “not knowledge independent of this or that experience, but knowledge absolutely independent of any experience”

³³ Before the term was Kant’s it belonged to the scholastics, who used it to capture a distinction found in Aristotle’s *Posterior Analytics*. According to Aristotle, a proposition A is prior to B in knowledge if it is possible to know A without knowing B, but not vice versa. The idea of a priori knowledge as something contrasted with empirical knowledge first appears in Leibniz, who defined empirical knowledge as knowledge of appearances and a priori knowledge as knowledge of the causes of experiences, which is gained through logical proof. (See 1703/1949.)

(CPR B3). A priori concepts are introduced in the same passage, where Kant adds that a piece of a priori knowledge is *pure* when it contains no concepts that can be derived only from experience. Unlike the term 'innate', however, 'a priori' cannot be applied to abilities. When Kant defined apriority as independence from experience, he didn't mean that the *cause* of knowledge should be independent of experience. He meant that the *justification* for that knowledge should be independent of experience. But abilities don't have justifications in this sense. A justification is what makes a belief into knowledge. But as Gilbert Ryle pointed out, we don't speak of believing-how, we only speak of knowing-how (Ryle 1949, 28). Thus, it makes little sense to talk about the justification for a piece of know-how. Apriority also gets applied to some things that innateness does not. However, most of these are variations on either a priori knowledge or a priori concepts that we do not need to concern ourselves with. For instance, Kant spoke of a priori intuitions. But the distinction between intuitions and concepts is really only a concern for those who embrace his particular architectonic. Gottlob Frege preferred to talk about a priori truths, rather than a priori knowledge. However it is easy to reduce the idea of an a priori truth to the idea of a priori knowledge. An a priori truth is a proposition that a normal human can know a priori. A final asymmetry between terms 'innate' and 'a priori' regards their application to concepts. While it made sense to define innate concepts separately from innate knowledge, it is easier to reduce a priori concepts to a priori knowledge. A concept is a priori if one can know a priori whether it applies to a given object a reasonable percentage of the time.

Kant's description of a priori knowledge has served roughly the same role in the discussion of apriority that talk of triggering has played in discussions of innateness: it is a simple definition adequate for most purposes. In the last section I described the notion of triggering as "not entirely vague," and I would certainly say no worse of Kant's definition of a priori. But just as we needed to expand on the idea of innateness in the last section, we need to expand on the idea of apriority in this section. The problem with Kant's definition is the ambiguity of the term 'experience'. There is a straightforward sense in which most of the promising conduits for a priori knowledge and concepts are not independent of experience. Proponents of a priori knowledge often say that it involves grasping truths by an intuition of some sort. But such intuitions seem on the surface to be experiences. Moreover, it will not do to simply restrict the scope of 'experience' to sensory experience, since we will want to allow knowledge and concepts acquired by memory to count as a posteriori. You might try to say that all a posteriori knowledge comes through the senses, and that memory is a posteriori knowledge because it requires justification, and that justification can only be sensory. But the fact that we remember something ought to carry some justificatory weight of its own, and this justification strikes me as a posteriori justification. Introspection, too, seems to be a posteriori, seems to have justificatory weight, but would be a priori if we defined a posteriori knowledge as sensory knowledge.³⁴

³⁴ I say that introspection *seems to be* a posteriori because it is sometimes hard to tell the

This section will survey the responses to this problem. Responses to this problem fall into two categories, which I will call positive and negative. The former response leads to the definition of ‘a priori’ used in this chapter. The latter response leads to a weaker definition of ‘a priori’, which will be used in the next section, where the definitions of innateness and apriority used are co-extensive. The positive approach to the a priori is the older one. It attempts to characterize a faculty of a priori intuition, and then define a priori knowledge and concepts in terms of this intuition. The two parts of the positive, or intuition-based, approach do not always go hand in hand. Detailed accounts of the nature of intuition have been proposed by many illustrious writers, from Kant to Husserl to Gödel to Charles Parsons. But the actual idea that a priority could be defined in terms of intuition has come from a different set of writers, generally working in the positivist tradition. Lawrence Bonjour’s recent book actually offers a detailed account of intuition in conjunction with a definition of apriority that falls into the opposing, negative camp. The negative approach to apriority simply characterizes experience and says that the justification of a priori knowledge must not appeal to it in some specific way. In addition to Bonjour’s work, definitions offered by Philip Kitcher and Tamara Horowitz fall into this category. As it turns out, the negative definition of a priori has a modal structure that makes it co-extensive with the dispositionalist definition of innateness. As a result, it will be used to argue for a combined nativist/apriorist position at the end of this chapter. In

difference between introspection and intuition.

fact, the two best negative definitions of apriority, Kitcher's and Bonjour's, are also co-extensive, so there will be no need to choose between them.

An important pitfall to avoid in this discussion will be to conflate the concept of apriority with other related concepts. At some time or another, apriority has been confused with necessity, universality, analyticity, unrevisability, and constitutivity. The root of this problem is Kant himself, who believed that all of these concepts save analyticity were co-extensive with apriority. After the positivists added analyticity to the list of co-extensive concepts, it became nearly impossible to keep them separate. However it is important that we do separate them. Necessity, as Kripke has emphasized, is a metaphysical concept, which applies to propositions if they are true in any possible world, whereas apriority is an epistemological concept. While it may turn out that all and only propositions that are known a priori are true in every possible world, there is no reason to assume this is true at the outset of our investigation. Similar remarks are true of universality, which is also a metaphysical concept. Analyticity is a linguistic concept, relating to the meanings of the terms used in a proposition. But there is no reason to assume at the outset of our investigation that ideas that are known a certain way can only be expressed a certain way. Unrevisability and constitutivity, on the other hand, are epistemological concepts. However, there is still no reason to assume at the outset that they are identical to apriority, or even co-extensive with it. A piece of knowledge is unrevisable if it could never be lost once it is possessed. While it may be that knowledge gleaned through certain channels has this remarkable property, it would be a bad idea to use it as a criterion for a priori knowledge, if only because it is so hard to exemplify. The

concept of constitutivity played a big role in the earliest discussions of apriority, even though it is now out of favor. In general, a piece of knowledge is constitutive of another piece of knowledge if it plays a role in the acquisition or justification of that knowledge. For Kant, this was the most important function of a priori knowledge. It continued to be so for the positivists, insofar as one needs to know the rules of a language before one can go on to use that language to describe things. Again, however, it would be hasty to assume that constitutivity and apriority are co-extensive. The human mind would be remarkably simple if all knowledge acquired through one channel was constitutive of knowledge coming through another. But the mind is never simple.

The positive approach to a priori knowledge attempts to identify a process of acquiring knowledge called an intuition, and then define a priori knowledge in terms of that process. To get a clear sense of what this intuition is supposed to be, it is best to glance briefly at the whole tradition of modern writing on intuition, not just those who use it to define a priori knowledge and concepts. Intuitions are generally defined as *immediate* sources of knowledge and concepts. They should require no outside help to bring us information. Because sensory knowledge is immediate, it is often classified as a kind of intuition. For clarity's sake, however, I will only use the term 'intuition' to refer to nonsensory intuition. There are two important distinctions among kinds of intuitions. First, Charles Parsons (1979–80/1996) has distinguished 'intuitions-of', which are ways of immediately grasping objects, and 'intuitions-that', which are ways of immediately

grasping propositions. The former kind of intuition is meant to resemble ordinary perception more closely than the latter.³⁵ There is also a distinction for specifically mathematical intuitions between those intuitions that are designed to secure the validity of inferences and those that are supposed to establish the truth of axioms.

Different theorists of the intuition have balanced these distinctions in different ways. Kant characterized intuitions (*anschauung*) using two criteria. The *immediacy condition* asserts that intuitions are representations (*darstellung*) formed by immediate contact with the thing they represent. The *particularity condition*, on the other hand, asserts that intuitions have to be representations of particulars (*CPR* B376–77 ff.).³⁶ The particularity condition implies that intuitions will always be intuitions-of for Kant. There is, properly speaking, no intuitive knowledge for Kant, because knowledge always requires both intuitions and concepts. Certain kinds of knowledge, however, take intuitions as their object in a way that makes them unique (*CPR* B33 ff.). One intuition Kant believes we possess is the *formal intuition*, an immediate access to the structure of our perceptions. This awareness is an intuition because our perception is structured by the forms of space and time, which are both particulars. Bodies of knowledge such as arithmetic, geometry,

³⁵ Intuitions-of should not be confused with intuitive concepts. Intuitions-of can have objects other than concepts, and intuitive concepts can arise from intuitions-that.

³⁶ There is some debate about the relative priority of the immediacy and particularity conditions. Hintikka (1969) argues that only the particularity condition is used to define intuitions and that Kant deduces the immediacy of intuitions from the particularity condition. Parsons (1969/1983) and Thompson (1972/1992), on the other hand, claim that immediacy is an independent part of the definition of intuition for Kant.

and algebra come from the application of general concepts to the formal intuition. The intuition is necessary both to provide axioms, where appropriate, and to secure inferences.³⁷ Parsons (1969/1983) has argued that intuition is specifically required to supply existence postulates. Existence cannot be established by concepts alone for Kant: there are things that are conceptually possible that are not actually possible. But geometry and arithmetic require certain things to exist in order to function. Geometry requires Euclid's first three postulates, which allow for the construction of the lines and circles. This is provided by the production of figures in the imagination. Although for Kant arithmetic does not have axioms, we still need some assurance of the infinity of the number line, and this can be provided by the intuition of the flow of time.³⁸ The role of the intuition in securing inferences is more general. Friedman, drawing on passages where Kant describes placing images before the mind in order to calculate with them, describes the intuition used to secure inference as the intuition used in "checking proofs step by step to see that each rule has been correctly applied, in short, the intuition involved in 'operating a calculus'" (1992, 92).

³⁷ Commentators typically claim that Kant's intuition only served one or the other of these functions. Lewis White Beck (1955/1965) and Parsons (1969/1983) claim that it only established the truth of axioms, while Friedman (1992) claims that it only secures inference. Here I am assuming that both sides are partially correct.

³⁸ Kant also felt that an intuition of space was involved in arithmetic, but Parsons minimizes its role: "I should venture to say that space enters the picture only through the general manner in which the inner sense, and thus time, depends on the outer sense, and thus space" (1969/1983, 62).

Edmund Husserl used the term ‘intuition’ to refer to a variety of forms of knowledge by acquaintance—sensory perception, phenomenological reflection, empathy, and categorical intuition. All of these are in some way intuitions-of. They involve the immediate grasp of an object. The last item on the list is the one that is of interest to us. The account of these intuitions we will look at is the one Husserl gives in volume II of the *Logical Investigations* §§40–47 (1900–1/1970). In the language of the *Logical Investigations*, knowledge is a matter of a mental act finding its “fulfillment” in an intuition. Roughly, this means that one finds the world the way one thought it would be. However, if thoughts are propositionally structured, it is clear that they have many parts that are not found in sensory intuition. Even observation sentences feature “little words” like ‘is’, ‘the’, and ‘and’, which do not correspond to perceptual objects. Husserl says these constitute the *form* of the sentence, whereas the elements of a sentence that can be found in sensation are its *matter*. Judgements that subsume one concept under another, such as “Red is a color” and other general statements, such as the axioms of arithmetic, are even farther from finding their fulfillment in sensory perception. Husserl’s solution is to say that in addition to the sensory intuition, perception involves a formal intuition that corresponds to the missing parts of the proposition. In this intuition, grasp of something nonperceptual is *founded* on the perception of a particular. For instance, the perception of a red object founds intuitions of the concepts ‘red’ and ‘color’. These intuitions are additional mental acts, applied to the acts that they are founded on. The objects of formal intuition are ideal, rather than real, and they are synthesized in the act of intuiting, rather than simply appearing as is.

Parsons takes Husserl's idea that intuitions are founded on particulars and weds it to Kant's idea of a formal intuition (1979–80/1996, 102). Parsons notes that when we see a word on a page or hear it spoken, we immediately perceive its type, while our awareness of it as a token is more peripheral. Our ability to immediately reidentify the type is evidence for this. "Typically," Parsons writes, "the hearer of an utterance has a more explicit conception of *what was uttered* than he has for the objective identification of the event of the utterance." He adds, "I believe that the same is true of other kinds of universals, such as sense qualities and shapes" (1979–80/1996, 102). These situations are everyday examples of intuitions of abstract entities founded, in Husserl's sense, on perception of particulars. Parsons imagines that the perception of at least some mathematical tokens works the same way, and gives the example of the perception of token in a simple mathematical language of horizontal strokes. (He explores this language more deeply in his essay "Ontology and Mathematics" [1971/1983].) The intuition of these types is an intuition-of, but is quickly leads to intuitions-that. For instance, we see "on the basis of a single intuition" (1979–80/1996, 105) that ||| is the successor to ||, because we perceive that the latter token is a part of the former, and with this perception, intuit that this property applies to all tokens of the same type. We can also intuit general propositions about indefinitely many types by relying on imaginations of tokenings. For instance, we can intuit that the stroke series can be continued indefinitely by imagining an arbitrary token string and seeing that we can add one to it. How do we imagine an arbitrary string of strokes? This is where Kant comes in. Our perception has a certain general structure, a form of intuition. When we imagine an arbitrary string of strokes, we

are using only this general structure. We are, in fact, exercising a formal intuition. The existence of the form of the intuition also gives us a way to explain the ‘may’ in the proposition “the stroke series may be extended indefinitely.” We are obviously not talking about physical possibility here. Instead, Parsons suggests, we are talking about what the form of our intuition would allow. In his essay, Parsons is mainly concerned with intuitions that propositions in his stroke language are true. However, Parsons’ intuition can also serve to secure the validity of inference. Recognizing the type of a token is a matter of being able to reidentify the type. This ability can reasonably be extended to cover recognizing instances of a schema. This should be all we need to operate a calculus.

Larry BonJour’s account of a priori intuitions is unique in that it eschews all analogies between intuitions and sensory perceptions, and with them all ideas of intuitions-of. In his recent book *In Defense of Pure Reason*, BonJour defines an intuition as an immediate grasp of the fact that a proposition is necessary. One knows a proposition a priori if one believes it on the basis of such an intuitive grasp. Although BonJour does not emphasize this, one could also say that one possesses a concept a priori if one’s application of it is guided by a proposition whose necessity is grasped intuitively (1998, 106). The grasp of the necessity of a proposition should not be understood by analogy with perception, however. Instead, BonJour thinks we should view intuition as arising out of the nature of thought in general. BonJour has a couple of reasons for shying away from the perception metaphor. First, if a priori intuition were a kind of perception, it ought to be associated with a particular organ or faculty, and there doesn’t seem to be one for mathematics. But

if intuition is just the product of thought in general, then the organ in question can simply be the brain (ibid., 109). Second, BonJour believes that something like Paul Benacerraf's famous argument against Platonism in mathematics could be applied to intuitions conceived of as perceptions. Benacerraf argued that if numbers were Platonic objects, we could have no knowledge of them, because knowledge requires causal interaction with the known, and Platonic objects do not interact causally. Similarly, if intuitions were modeled on perception, they must be intuitions of particular objects, presumably Platonic concepts and propositions. But perception is a causal process, so intuitions as perceptions are self-contradictory. To explain how intuition can arise out of the workings of thought in general, BonJour proposes we return to an Aristotelian and Thomistic conception of the mind (ibid., chap. 6.7). In BonJour's view, Aristotle and Aquinas were right to think that abstract entities actually enter the mind when we contemplate them. This contact is similar to the contact between abstract entities and the material objects that instantiate them, but the two forms of contact are not quite the same. When I think about triangles, my mind possesses the form of triangularity. It does not possess the form as material objects do, for then it would become literally triangular. My mind has the *esse intentionale* of triangles, rather than the *esse naturale*. BonJour breaks from Aquinas and Aristotle, however, over how to distinguish the two forms of contact with abstract entities. Aquinas and Aristotle seem to have believed that there is one form of triangularity that can be possessed in two ways. BonJour instead believes that there are two forms of triangularity, which are united by an über-form. In either case, because our minds actually possess abstract entities, it is easy to grasp their necessary properties.

When I have the rational intuition that nothing can be red and green all over at the same time, my mind possesses a complex universal, of which redness and greenness are literal constituents. I am aware of the necessary features of redness and greenness because I am aware of the state my mind is in.

None of the above writers used the notion of an intuition to actually define a priori knowledge. Most, in fact, presuppose a notion of apriority in outlining the notion of an intuition. Other writers, however, have drawn on this tradition of thinking about intuitions to try to derive a definition of apriority. An article by David Benfield is a good example of approach. Benfield makes the following equation: “S knows *p* intuitively = Using only ordinary human capabilities S knows *p* directly and not in virtue of any other knowledge S has” (1974, 152). Sensory knowledge of the external world is not a priori under this definition, according to Benfield, because we need further evidence before we can be sure that our senses are not deceiving us. Although this definition is a little thin, we can substitute any of the more sophisticated accounts of intuition seen in the last paragraph. Exactly which one is best to use will be decided when we get to the arguments for intuition-based apriorism—we need to see what kind of intuition these arguments will support. Given a definition of intuition, Benfield goes on to define a priori knowledge as intuitive knowledge, along with any knowledge that can be deduced from intuitive knowledge using intuitively valid principles (*ibid.*). Although Benfield’s article is fairly recent, his basic style of definition goes back to the positivist era. Albert Casullo (1977), commenting on Benfield’s article, offers a similar account of a priori knowledge and cites

as his predecessors a remark by C. D. Broad (1936, 103) and a definition by Ambrose and Lazerowitz (1948, 17), which was endorsed by Arthur Pap (1958, 95).

If we define a priori knowledge as knowledge that comes to us by intuition, the task of the mathematical apriorist is to show that there is such a thing as an intuition, and specifically a mathematical intuition. I will now look at four arguments to this effect. The first will be an argument for a specifically mathematical intuition, while the remaining three, drawn from the work of Larry BonJour, will be for the existence of a priori intuitions in general.

The most common motivation for belief in a mathematical intuition is that it gives mathematics a foundation apart from the empirical sciences. The typical foil for this kind of argument is Quine, who believes that mathematics consists of very high level generalizations about the empirical world. There are three main arguments that say that mathematics should have a foundation outside of the empirical sciences and that this foundation should take the form of a mathematical intuition. The first is an appeal to the *obviousness* of mathematical knowledge. Many statements of mathematics are universally regarded as obvious. But, as Charles Parsons points out, if mathematics were a part of the empirical science, these statements would be “bold hypotheses, about which a prudent scientist would maintain reserve, keeping in mind that experience might not bear them out” (1979–80/1996, 102). Penelope Maddy puts the point more starkly, “Isn’t it odd, to think of ‘ $2 + 2 = 4$ ’ or ‘the union of the set of even numbers with the set of odd numbers is the set of all numbers’ as highly theoretical principles?” (1990, 31). The very obviousness of some mathematical statements indicates that they are not parts of an

abstract physical theory, but are rather known immediately, which in turn indicates that an intuition is at work. Maddy adds to this argument an appeal to the *autonomy* of mathematics. “Many of us tend to think of mathematics, not as a highly theoretical adjunct to physical science, but as a science in its own right, with its own subject matter and its own methods” (ibid., 45). Elsewhere she notes, “Mathematicians have a whole range of justificatory practices of their own, ranging from proofs and intuitive evidence to plausibility arguments and defenses in terms of consequences” (ibid., 31). But if mathematics is an autonomous science, then the epistemological statuses of its propositions should be parallel to those of scientific statements, with some propositions having the status of immediate, foundational beliefs, and others theoretical propositions. “Some mathematical beliefs should be basic and non-inferential, just as some scientific beliefs are” (ibid., 46). The final argument of this sort is that if mathematics did not have a foundation apart from empirical science, the huge swaths of mathematical belief that are never applied would be completely unjustified. This despite the fact that mathematicians engage in elaborate procedures such as deductive proof that they think will justify the statements of pure mathematics. “Mathematicians,” Maddy notes, “are not apt to think that the justification for their claims waits on the activities in the physics labs” (ibid.).

To these arguments for a specifically mathematical intuition, we can add three arguments from Lawrence BonJour’s recent book *In Defense of Pure Reason* for the need for a priori intuitions in general. All three are based on the need to reply to a skeptical challenge. The first two are given very briefly in the opening chapter of the book (1998, 3 ff., 5 ff.). There BonJour claims that some sort of a priori justification is necessary if we

are to either have knowledge of the empirical world or ever make a justified inference. Empirical knowledge requires a component of a priori justification because it must contain something that goes beyond the sensory given. The given, for BonJour, is the set of “foundational” beliefs that are fully justified by experience alone. According to BonJour, his argument makes no assumption about whether these foundational beliefs deal with objects or private sense data. No matter what the given consists in, the move beyond it must be justified. But if this justification is to avoid circularity, it must come from outside the sensory given. Hence it must be a priori. The need for a priori justification in inference is brought out by *reductio*. If inferences were justified empirically, then their empirical justification could be written in as a new premise. The resulting set of premises will either contain the conclusion explicitly or not. “In the former case, no argument or inference is necessary, while in the latter case, the needed inference clearly goes beyond what can be derived empirically” (*ibid.*, 5). In both arguments, the alternative to a priori justification is a kind of radical skepticism, a solipsism of the present moment in the first case and a denial of the validity of all inferences in the second. The final step of each argument is to say that the requisite a priori justification is best viewed as an a priori intuition.

BonJour’s third argument, which takes up the second and third chapters of his book, claims that the two alternatives he sees to an intuition-based apriorism fail. One can challenge intuition-based apriorism either by denying the existence of a priori knowledge altogether, or by denying that it comes from a mysterious faculty like an a priori intuition. ‘Moderate empiricism’ is BonJour’s term for the positivist viewpoint that grants the

existence of a priori knowledge, but attempts to rule out spooky intuitions by saying that a priori knowledge is limited to analytic statements. Moderate empiricism fails, according to BonJour, because there is no notion of analyticity that leads to a conception of justification that is any more clear than a priori intuitions. 'Radical empiricism' is BonJour's term for Quine's doctrine that there is no a priori knowledge at all. BonJour has several objections to radical empiricism, but the only one I will talk about here is his claim that the radical empiricist has no reply to the skeptic. According to BonJour, Quine cannot show (1) why any of his criteria for good beliefs are justified, (2) why any application of his criteria is justified, and (3) why any belief should ever be revised. On the first count, although Quine offers us many standards by which we may judge a belief—simplicity, fecundity, conservatism—he can do nothing to show that they are truth conducive. An apriorist could give a transcendental deduction; Quine can't. Quine's only alternative is to appeal to his own standards to justify his own standards, which is question begging. On the second count, Quine cannot ever explain how his criteria for belief should be applied. Any judgement that a certain criterion has been applied correctly is itself up for judgement, and this regress can only be stopped by an a priori justification. On the third count, Quine can never explain why we should ever revise our beliefs. Supposedly we revise our beliefs because of recalcitrant experience. But the claim that any experience is recalcitrant is itself a claim that can be added to the web of belief. The claim that the experience is recalcitrant with regard to this expanded web of belief is another sentence that must be added to the web of belief. An infinite regress ensues, and this regress must be traversed before any claim is revised.

There are problems with all three of Bonjour's arguments, but only the difficulty with the first argument is fatal. Bonjour's first argument said that a priori justification was necessary to go beyond the given in empirical knowledge. There are many complaints one can make against the distinction between 'the given' and 'the inferred', but for now, assume that we can draw such a distinction. Despite his assertion to the contrary, Bonjour's first argument makes strong assumptions about the nature of the sensory given. If we require a priori justification to move beyond the given, then the given must have no conceptual articulation at all. It would have to contain only the phenomenal "red here now." Few empiricists would accept such an old-fashioned, phenomenalist notion. Anyone who wants to isolate perception as a source of knowledge today would want to claim that when we perceive, we perceive *objects*, and that perceiving objects does not involve deploying a priori concepts. But once Bonjour grants the perception of objects, it is possible to talk about making inferences beyond the given without appeal to the a priori. Objects can be reidentified, correlated, and manipulated in ways that generate knowledge, solely in virtue of being objects. Bonjour might reply that this kind of knowledge involves an a priori concept, namely, the concept of object. But if this is all he means by a priori, then I think most empiricists would grant it to him. A modern empiricist wants to say that nothing is required for knowledge that can't be found in the perception of objects. If you want to label some portion of the perception of objects a priori, that is your own terminological vice.

Bonjour's second and third arguments are both demands for kinds of metajustifications, arguments that show why another argument or form of argument is

valid. Bonjour's second argument demands a metajustification for ordinary inference—a reason why we should believe that the conclusion is true if the premises are true. Bonjour's third argument demands that Quine produce a metajustification of his basic epistemic principles, such as simplicity and conservatism. The legitimacy of both these arguments is impugned by Bonjour's unwillingness to provide an equivalent metajustification for his own favored epistemic tool, the rational intuition. Just as Bonjour asked the empiricist in general to justify ordinary inference, and Quine specifically to justify his epistemic principles, we can ask Bonjour to explain to us why we should trust the deliverances of intuition. All of the aspects of the demands Bonjour made on the empiricists would carry over into the demand that we make on Bonjour. When Bonjour demanded that the empiricist explain why an ordinary inference is valid, he assumed that this metajustification would be a part of the inference it justified. As an empirical statement, he said, it could be added to the list of premises in the argument. We can say the same about the demand for a metajustification of the a priori intuition. By the standards imposed by Bonjour's second argument, every time one relies on an intuition, one must implicitly invoke a metajustification of intuition as a source of belief. Both Bonjour's request for a metajustification of ordinary inference and his request that Quine produce a metajustification of his epistemic principles assumed that metajustifications were necessary to reply to the skeptic. We can say the same about our demand for a justification of intuition: if one is not produced, the specter of skepticism will materialize and scare off all of our knowledge claims. Bonjour considers this kind of challenge in section 5.5 of his book, although he does not make the connection between the demand

for metajustification against him and the demands he made of the empiricists. In this section he seriously undercuts much of the rest of his book by claiming that he does not need to provide a metajustification for rational intuitions. Bonjour offers two reasons for this. First, the task is simply too hard, especially if the metajustification is supposed to be a part of the original a priori justification. If the metajustification is supposed to be a part of the original inference, it would have to itself be a priori. But it is extremely unlikely that such a proof could be given: “That beings like ourselves have rational insights that are even generally correct, or, weaker still, correct more often than not...does not appear to be even an initially plausible candidate for the status of metaphysical necessity” (ibid., 144). But if an a priori metajustification is both required and impossible, then a priori knowledge is *prima facie* incoherent, and the empiricist wins too easily: “Even a convinced radical empiricist should, I think, be dissatisfied with a victory that comes as cheaply and as easily as this. While we can perhaps understand how the idea of a priori justification might turn out to be incoherent in some relatively deep or complicated way, it seems difficult or impossible to believe that there is not even a *prima facie* coherent concept that generations of rationalists and moderate empiricists have had in mind” (ibid., 145). Second, demanding that the rationalist produce a metajustification assumes that intuitions are not legitimate sources of knowledge in themselves, “but instead function merely as a kind of earmark or symptom for picking out a class of believed propositions that the supposedly required metajustificatory premise then tells us are, on some independent ground, likely to be true” (ibid.). This assumption, Bonjour claims, is question begging.

I actually think that the demand for metajustification is unreasonable, for reasons that will come out in the next chapter. For now, my interest is in rescuing BonJour's argument, and the best way to do that is to drop the claim that rational intuition needs no metajustification. Certainly his attempts to avoid metajustification are the weaker of the arguments in conflict. Saying that an objection is too successful is never an effective reply. Moreover, if we are taking skepticism seriously, then one does not need to assume that intuitions are mere symptoms or earmarks of true propositions in order to demand a metajustification. Metajustifications are always needed to reply to the skeptic. Of course, if we say that intuitive knowledge requires a metajustification, we will actually have to produce such a metajustification. But given our dialectical situation, this is actually not as hard as BonJour thinks. What we need to show is that *if* there is a faculty of intuition that matches our description, its deliverances would meet some standard of reliability. Since our goal here is to show that the apriorist's own theory is able to meet the demands placed on others' theories, we are able to assume that there is an intuition that meets our description—by being an access to the necessary structure of our perception, or by being a grasp on the forms that one's mind participates in, for instance. Nor is it necessary to argue that intuitions are infallible. The arguments for a specifically mathematical intuition pointed to a faculty that was spontaneous and could allow us to conceive of mathematics as a discipline independent of empirical science. This faculty does not need to be infallible any more than its counterpart in the empirical sciences, ordinary perception, does. The one argument from BonJour that we are accepting that might make us think that intuition must be infallible is the argument that claims that intuitions are necessary to

provide a metajustification for ordinary inference. But such a metajustification would only have to be infallible if ordinary inference was infallible, which clearly it is not.

There is one condition on our metajustification of intuition that will limit our options. It must be easier for the apriorist to show that intuition as she has conceived it is reliable than it is for someone like Quine to show that his epistemic standards are reliable. For this reason we cannot accept the account of rational intuition BonJour gives with his arguments. BonJour, you will recall, claimed that intuitions arose because the mind participated in the forms it contemplated. Given the extensive metaphysical baggage this comes with, an empiricist could too easily reply to our metajustification by saying “Sure, if we believe in the forms, and if we believe that our mind becomes a part of them when we think, then it is entirely reasonable to believe that we have an immediate grasp of the necessity of necessary propositions. But how plausible is the antecedent to this conditional? No matter how strained the empiricist metajustifications of ordinary inference and our basic epistemological principles were, they have to be more likely than this.” Fortunately, none of the arguments produced so far have wedded us to BonJour’s version of the intuition. I propose instead that we adopt Parson’s hybrid Kantian/Husserlian conception. Parsons avoids making excessive claims about the mind’s ability to grasp pure ideas by founding his intuition on ordinary perception, as Husserl did. Rather than talking about the mind participating in forms, he gives us an everyday example of perceiving a token and having immediate access to its type. Parson only asks us to believe two things: there is a structure to our perception, and we can access this structure though intuitions of types founded on perceptions of tokens.

Given this arrangement, the metajustificatory question becomes, “How do we know that our intuitions about the structure of our perceptions meet some minimal standard of reliability? For instance, how do we know that they are right more often than not?” This line of questioning can be buttressed by cases where we have been wrong about the structure of our perception, such as the rise of non-Euclidean geometry. The answer is appropriately Kantian: If our intuitions are not right more often than not, experience would not be possible. The argument will only deal with mathematical intuitions, and will proceed by cases. As we noted earlier, mathematical intuitions are used to provide axioms and to secure the validity of inferences. The intuition we have argued for needs to be able to fill both these roles. On the one hand, we simply argued that intuitions were needed to secure inference, so our intuition had better be able to do this task. On the other hand, we argued that intuitions were needed to provide mathematics with a basis apart from empirical science. Intuitively known axioms would be a good way to do that. These two uses of intuition, to establish axioms and to secure inferences, will be our two cases. In each case, if intuition is not giving us an accurate picture of the structure of our perception more times than not, our perception simply could not exist in the way we know it to exist. First consider the role of intuitions in securing inference. We described this as the intuition necessary to operate a calculus. Now suppose that our intuitions about how to operate calculi were incorrect most of the time. In such a case, calculi could not exist because there would be no correct answers to any problems posed in them. But calculi with discernable right answers are a part of experience as we know it. Therefore our intuitions about how to operate a calculus must be right more often than not. Next

consider the role of intuitions in providing us with axioms for mathematics. Here the intuition functioned by producing models in the imagination. In this case, the conclusion of the metajustification is stronger than the previous case. Our intuitions here actually cannot be incorrect. If we have produced something in our imagination, we have produced it in our imagination. In a way, these existence postulates are like self-fulfilling prophecies.

The Combined Nativist-Apriorist Position

The two viewpoints I have examined so far depend on the existence of certain kinds of entities. The representationalist version of nativism required us to believe that there are innate representations, while the intuitionist version of apriorism required us to believe in a mathematical intuition. There are weaker definitions of both ‘innate’ and ‘a priori’, and these two dovetail nicely, leading to a combined nativist-apriorist view of mathematical knowledge. The task of this last section of the chapter is to outline these two definitions, show that they are co-extensive, and then examine the evidence that mathematical knowledge is innate and a priori in these senses.

We have been classifying definitions of innateness based on how they separate learning from triggering. The nativist needs a way to determine when the content of a capacity lies in the individual and when it lies in the events that cause that capacity to be expressed. The representationalist nativist did this by locating a representation in the mind of the individual. In addition to tying the existence of innate knowledge to the existence of a particular entity, the representationalist rules out many routes to innate

knowledge that others accept, such as knowledge acquired by timing and architectural mechanisms. A more catholic approach to innate knowledge distinguishes learning from triggering via the notion of a disposition. This is actually one of the oldest definitions of innateness. When René Descartes introduced the doctrine of innate ideas to the modern debate, he described them as a sort of disposition (1647/1973, 442), and he compared them to a disposition to a certain illness. Just as a child who inherits a disease may not be symptomatic at birth, but already has within her the propensity that will cause her to become ill later in life, so too we may be born predisposed to acquire certain sorts of knowledge. This passage from Descartes has been picked up in contemporary times by Steven Stich, who has used it in the introduction to the collection *Innate Ideas* to formulate a rigorous definition of innateness in terms of dispositions. He formulates the definition in terms of beliefs, which he describes as either being occurrent or dispositional, depending on whether they are being consciously contemplated: “A person has a belief innately at time t , if, and only if, from the beginning of his life to t it has been true of him that if he is or were of the appropriate age (or at the appropriate stage of life) then he has, or in the normal course of events would have, the belief occurrently or dispositionally.” (1975, 8). Innate beliefs are thus a complex counterfactual. A person has an innate belief if at one time it is true of them that had it been another time later in their life, they would have the belief in question, given the typical course of events. Since beliefs not actively contemplated are already a sort of disposition, an innate belief may be a disposition to a disposition. The two levels of disposition do not collapse into each other, however. The disposition possessed by someone who holds a belief dispositionally

is a disposition to hold that belief consciously *as soon as the issue is broached*, whereas the disposition possessed by someone with an innate belief is a disposition to hold the belief after a period of experience.

In this definition, of course, much rides on what is to count as “the normal course of events.” Here Stich considers two possibilities, either the “normal course of events” means having enough experience to possess the concepts involved in the belief in question, or it means having enough experience to possess concepts at all. If the innateness of mathematical beliefs is to be a live issue at all, we will have to adopt the former option. It is entirely possible to be able to have concepts but hold no mathematical beliefs. On the other hand, it is reasonable to think that as soon as one has had enough experience to possess the concept of number, one has basic arithmetic beliefs. It is also plausible to say that beliefs about higher realms of mathematics are held as soon as one possesses the concepts involved in them. Thus when one truly understands transfinite cardinals, one possesses beliefs about the Lowenheim-Skolem theorem.³⁹ The issue of the innateness of mathematical knowledge then becomes tied to the truth of this view.

Stich does not address the question of how to transform this analysis of innate belief into an analysis of innate knowledge, but the form such an analysis must take is pretty

³⁹ Conversely, one would not understand transfinite cardinals *until* one possesses beliefs about the Lowenheim-Skolem theorem. This way of thinking about mathematical beliefs is another expression of the idea that the concept of a mathematical object exhausts its properties. There is nothing more to them than what we grasp when we are able to grasp them.

much forced on us. For an innate belief to be knowledge, it must at least be true and somehow justified or warranted. We must also say that, for an innate belief to be knowledge, the disposition for belief that is present at birth must also be a disposition to possess a justification. If we did not make this stipulation, then the *knowledge* of the proposition would not be innate. Suppose I believed innately in some proposition *p*, but had no tendency to justify that belief. Then my belief in *p* is only accidentally related to its truth. My knowledge would only come later, when I accessed the correct justification. But this would mean that my *knowledge* of *p* is not innate.⁴⁰ The knowledge came from the learning process that occurred after birth. These considerations add up to a nice analysis of innate knowledge. An innate belief is innate knowledge if and only if it is true, and the process that allows one to form that belief at the appropriate age under the normal course of events is a warranting or justificatory process. I should emphasize that under this analysis, the disposition that is present at birth is not what justifies the belief later on. Rather, we have a disposition at birth to acquire a belief later in life and to acquire a justification for that belief. Under this analysis, innate knowledge is dispositional in a way that ordinary knowledge is not. One might innately know a proposition without yet believing it occurrently or dispositionally. It is, however, counterfactually true of one that

⁴⁰ This is, admittedly, just the position traditional nativists thought we were in. For instance, Descartes, in the *Meditations*, first discovers that he has innate beliefs, and then decides that they are justified. This just goes to show that the versions of nativism I am discussing here are stronger than traditional nativisms like Descartes'.

if one were the right age, one would have the belief in question, occurrently or dispositionally. If this seems counterintuitive, we can always add the additional rider that for an innate belief to be innate knowledge, one must not only have had the requisite disposition at birth, one must now have the belief in question, either dispositionally or occurrently.

To extend this analysis to innate concepts we need a new understanding of “the normal course of experience.” It will obviously not do to say that it is “sufficient experience to understand the concepts involved.” Here I am going to appeal to the next more basic level of mental faculties: possession of abilities. A person has a concept innately if at birth it is true of them that if they had the abilities required to manifest their possession of that concept—such as the motor skills and short-term memory necessary to respond to stimuli appropriately—then they would manifest possession of that concept. Note that I am *not* saying that a concept should be considered innate if one would manifest possession of it if one possessed all of the skills necessary to use the concept. This would make all concepts trivially innate. If the nature of a concept is simply its correct use, then it is true of any concept that if one had the skills involved in its use, one would manifest possession of the concept. Suppose we said that the concept of number was constituted in part by the ability to count. We do not want to say that one possesses a concept of number innately if one would manifest that concept as soon as one is able to count. This would make number trivially innate. Rather than saying that a concept is innate if it is manifested when one has the abilities which constitute that concept, I am saying that a concept is innate if it would be manifested as soon as one had the abilities

necessary to use the abilities that constitute a concept. Counting requires an ability to point and to put collections in one-to-one correspondence. A concept of number would be innate if one manifests it as soon as one possesses these background abilities. We reach rock bottom when we talk about innate abilities. There is no more basic concept to use to define the relevant course of experience. So in defining innate abilities, I will give up on talk of triggering altogether. Innate abilities are those abilities that are manifest at birth.

The dispositional definition of innateness is quite weak compared to the representationalist definition. It is making no claim about the nature of the structures that give rise to innate capacities; it only claims that from birth they have a tendency to come into play in the right circumstances. A correspondingly weak definition can be offered for a priori capacities. When I introduced the intuition-based definition of apriority, I described two approaches to the concept, one positive and one negative. The intuition-based definition was the positive approach. It attempted to define a channel of knowledge that we can call “a priori.” The negative approach, in contrast, attempts to define “experience,” and then defines a priori knowledge as knowledge that does not depend on experience for its justification. A simple example of this approach is Tamara Horowitz’s essay “A Priori Truth” (1985). Horowitz stipulates that, “a person has enough experience to know a proposition if and only if that person’s experience suffices for him to discover that the proposition is true given only further ratiocination” (ibid., 230). She goes on to say that a person *a* knows a proposition *p* independently of experience “if and only if in every world in which *a* exists, *p* is true, and *a* has the concepts in *p*, *a* has enough experience to know *p*.” Furthermore, “A truth *p* is a priori if and only if everyone can

know p independently of experience.” In this definition, the idea of having enough experience to know something plays the role of ensuring that a priori justification be free of experience in the appropriate fashion. Unfortunately, Horowitz’s definition contains a fatal circularity. In specifying experience, she uses the notion of ratiocination, however there is no way to distinguish ratiocination from other experience without first assuming a notion of experience.

More promising negative definitions come from Larry BonJour and Phillip Kitcher. In his recent book *In Defense of Pure Reason* (1998), BonJour declares experience to be “any sort of process that is perceptual in the broad sense of (a) being a causally conditioned response to particular, contingent features of the world and (b) yielding doxastic states that have as their content putative information concerning such particular contingent features of the actual world as contrasted with other possible worlds” (ibid., 8). A proposition is then said to be justified independently of experience if no appeal to experience is needed to justify it once the proposition has been understood (ibid., 10). Although this definition obviously links apriority closely to metaphysical necessity, it does not conflate the two ideas. BonJour’s notion of apriority remains an epistemological concept, not a metaphysical one. In fact, BonJour’s apriority presupposes a notion of necessity, since we must be able to identify the contingent features of our universe to apply it. The other promising negative approach to a priori knowledge is the one we have already spent some time examining: Philip Kitcher’s account. Kitcher, you will recall, says that S knows P a priori if and only if in every possible world where P is true and S has the cognitive abilities distinctive of humans and enough experience to contemplate P ,

S also has a warrant of the same type available for believing that *P*. A warrant is a belief-forming process that is reliable or in some other way privileged as the “right way” to come into a belief. While this definition does not contain a direct definition of experience, it makes no effort to specify a channel for a priori knowledge. So it is more than reasonable to call this a negative approach.

BonJour’s definition is equivalent to Kitcher’s: BonJour would call a piece of knowledge a priori iff Kitcher would as well. One might think that it is impossible to compare the two because BonJour does not share Kitcher’s psychologistic framework. BonJour’s definition of experience does bring in the notion of causality, however, and this will give us enough common ground to make a comparison of the two definitions. First we will show that a piece of knowledge that meets BonJour’s definition of apriority also meets Kitcher’s; then we will show that the converse holds. Suppose that *S* knows *P* and does not need to appeal to a causally conditioned response to a particular, contingent feature of the world in order to justify her belief that *P* once she has understood *P*. We need to show that *S* would have what Kitcher would call a warrant available for *P* in all possible worlds where *S* has enough experience to contemplate *P*. But if whenever *S* understands *P* she needs no appeal to causally conditioned reactions to particular, contingent features of the world, then her justification must appeal to necessary or universal features of the world, or perhaps to causally conditioned reactions to necessary or universal features of the world. But these will be available to her in any possible world where she has enough experience to contemplate *P*. All that remains is to say that the warrants that are available in any of the relevant possible worlds belong to the same type.

I claim that in any decent typology being an appeal to necessary or universal features of the world or to causally conditioned responses to necessary or universal features of the world will be an important classification. Therefore, if BonJour says that *S* knows *P* a priori, Kitcher would also say that *S* knows *P* a priori.

To show the converse entailment, assume that *S* knows *P*, and has a warrant for *P* available in any possible world where she has enough experience to contemplate *P*. In nonpsychologistic terms, this would mean that she has a justification in any possible world for believing that *P*. The simplest way for this to be true would be if her justification depended only on universal or necessary features of the world, in which case her knowledge would be a priori by BonJour's definition. One might think that it is also possible that *S*'s knowledge is still dependent on causally conditioned responses to particular, contingent features of the world because in any possible world there is some *other* causally conditioned response to a contingent feature of the world she is dependent on. This possibility is ruled out, however, by Kitcher's stipulation that the a priori knower's warrant must belong to the same type in any possible world. Although no typology of warrants has ever been provided, I think it is safe to assume that warrants that depend on different contingent features of the world belong to different types.

Given that Kitcher's definition and BonJour's definition are equivalent, I don't think there is a need to choose between them. They are both different ways of formulating the same negative notion of apriority. The only task remaining is to extend this definition to cover a priori concepts and abilities. As I mentioned in the section on intuition-based apriorism, abilities cannot be a priori because we cannot talk of their justification. So the

only task is to define a priori concepts. Now another point that came up in the section on intuition-based apriorism is that it is easy to define a priori concepts in terms of a priori knowledge. A concept is a priori if one can know a priori whether it applies to an object a reasonable percentage of the time. The simplest way to translate this into negative apriorist terms would be to say that a person possesses a concept a priori iff in any possible world where she has the abilities typical of humans and enough experience to contemplate the concept, she knows whether that concept applies to an object. I would like to make a slight alteration in this, however. In the dispositional version of innate concepts, I specified that in the relevant range of possible worlds the person must have the abilities necessary to manifest possession of the concept. (This was because the equivalent clause in the definition of innate knowledge referred to sufficient experience to understand the concepts involved in the proposition known.) To keep our definitions parallel, I would like to say that a person possesses a concept a priori if in any possible world where she has the abilities typical of humans and sufficient abilities to apply the concept in question to an object, she would know whether the concept applied to an object a reasonable percentage of the time.

Now the interesting thing about both the dispositional definition of innateness and the negative definitions of apriority is that they involve complex counterfactuals. They talk about what happens in a range of possible worlds. This enables us to reach an unusual conclusion: under these definitions, innateness and apriority are co-extensive. Most philosophers take it as a piece of common sense that any findings that deal with innateness will have no impact on the issue of apriority. W. D. Hart has stated flatly, "For

the problem of a priori knowledge, innate ideas are an inane solution.” (1975, 109). His basic argument is that saying an idea is inborn is no guarantee that such an idea is true. It is quite easy for us to be programmed with false beliefs. (Hart does not provide an example, but they are not hard to come by—we might have an innate belief that space is Euclidean, when in fact it is curved in proportion to the mass it holds.) Because saying that an idea is inborn is no guarantee that it is true, appealing to the innateness of a belief can in no way justify it. Hart concludes that innateness and apriority must belong to entirely different realms of discourse: “Talk of innateness is appropriate in the context of discovery, yet the problem of a priori knowledge is posed wholly within contexts of justification” (ibid.). In the same volume, Steven Stich offers the converse point. Not only is innateness not sufficient (nor even helpful) for establishing apriority, it is not necessary. Beliefs may be caused by experience, but have justifications that are “not to be found in the experience that caused the beliefs, nor in any other experience” (1975, 17). Stich, like Hart, concludes that innateness and apriority belong to completely different domains of discourse. “To say that a bit of knowledge is a priori, then, is to say something about its justification, while to say that a belief is innate is to say something about its cause of genesis.” (ibid.).

If we adopt Stich’s own definition of innateness and Kitcher’s analysis of knowledge, specifically his definition of apriority, then Stich and Hart are wrong about the relation between innate and a priori knowledge. Hart is quite right to point out that it is possible to have a false innate belief, but *if an innate belief is knowledge*, then it is a priori knowledge. Conversely, if a piece of knowledge is a priori, it must be innate. This

happens because the two definitions share a modal structure. Stich's definition, you will recall, goes like this: "A person has a belief innately at time t , if, and only if, from the beginning of his life to t it has been true of him that if he is or were of the appropriate age (or at the appropriate stage of life) then he has, or in the normal course of events would have, the belief occurrently or dispositionally." (1975, 8). "The normal course of events" it turned out, meant having enough experience to understand the concepts involved. A person thus has an innate belief if a certain counterfactual holds true: if they had enough experience to understand the concepts involved in the belief, and if events had proceeded normally, they would have the belief in question. To turn this definition of innate belief into a definition of innate knowledge we added the stipulation that the belief is true and that the disposition includes a disposition to possess a justification for that belief. As I noted on page 200 of this dissertation, without this stipulation, the *knowledge* of the proposition would not be innate. While the disposition for the belief was present at birth, there would be no disposition to possess the justification of that belief.

Kitcher's definition of apriority also involved a counterfactual: in any possible world where S has enough experience to contemplate p , S has available a warrant for believing that p . In each case, we say that a belief has the property in question if in any course of events where one has enough experience to contemplate the relevant concepts, one will also be able to form the belief. A minor difference between the two is that Kitcher has a stricture on the range of possible worlds that Stich does not explicitly put on the "normal course of events." Kitcher only considers those possible worlds where one has the abilities characteristic of humans. Stich might implicitly be operating with such a

restriction, given that he is considering the possible course of events of the development of a human baby. At the very least, it is in the spirit of his analysis. In what follows, I will assume that the “normal course of events” means developing the characteristics distinctive of humans. The main difference between Kitcher’s a priori and Stich’s innate lies in the fact that for Kitcher, the process that is available in a range of possible worlds does not just generate belief. It is a warranting process. But as I will now show, this does not keep “innate” and “a priori” from being co-extensive.

To see that innate knowledge must be a priori, suppose I know a proposition p innately. Then, it was true of me at birth that were I to have a life long enough to contemplate the concepts involved in p , I would be able to know p . If my innate belief is to be innate knowledge, then I must also have a disposition to possess a justification or warrant for that belief at the appropriate age. Therefore in any possible world where I have enough experience to contemplate p I have a warrant for believing that p . The only portion of Kitcher’s definition left is that these warrants all be of the same type. This point is harder to argue, because no one has developed the relevant typology of warranting processes. However, I would argue that whatever typology one develops, it would classify all of the warrants involved in innate knowledge together. We know they have at least one interesting property in common!

The proof that a priori knowledge must be innate is quite brief. Suppose I know a proposition p a priori. In any possible world, then, where I have the life experiences needed to contemplate the concepts used in p , I will have a warrant to believe in p . Therefore it was true of me at birth that under the normal course of events, I would come

to believe that p , either occurrently or dispositionally. Furthermore, the process that would lead to this belief would be a warranting one. Therefore I know p innately.

One might think that these two short arguments beg the question by stipulating that the innate belief must be true and created by a warranting process. However this stipulation is forced on us by the fact that we are concerned with innate *knowledge*. It is clear that innate *beliefs* and a priori knowledge are not co-extensive, for all the reasons that Stich and Hart emphasize. If we are only talking about beliefs, there isn't a chance that there will be a link between innateness and apriority. Once we talk about innate knowledge, the possibility of a link between innateness and apriority becomes a live issue. Simply switching to talk of innate knowledge doesn't guarantee that innateness and apriority will be co-extensive, however. The right definitions need to be adopted. Innateness and apriority would not be co-extensive if we used an intuition-based definition of apriority with the dispositionalist analysis of innateness, for instance. We settled on Parsons' notion of a mathematical intuition, where we grasp a limitation of what the forms of space and time will allow through the perception of a particular. But nothing in the description of this kind of intuition implies that we are born with a disposition to have them. It is at least logically possible that these intuitions are learned. So the stipulation that innate beliefs must be true and justified is not enough to guarantee that innateness and apriority are co-extensive, but it is necessary to broach the issue. For this reason, I take it to be a legitimate assumption.

One might argue that the only definition that needs to be weak is the definition of apriority. Once we adopt the negative definition of apriority and assume that innate

beliefs are true and justified, the result is fixed. Innateness and apriority have to be co-extensive. Suppose this were the case. Suppose we adopted the representationalist definition of innateness and assumed that our inborn representations were true and justified. I concede that all innate knowledge would then be a priori on the negative definition of a priori. But not all a priori knowledge would necessarily be innate. There may be things that we have warrants to believe in the appropriate range of possible worlds that we have no innate representation of. If Kitcher is right when he suggests that “I exist” is a priori on his definition, then it would almost certainly be a case of knowledge that is a priori on the negative definition of a priori, but not innate on the representationalist definition of innate. So, again, it seems that it is necessary to adopt the weakest definitions of innateness and apriority in order to make them co-extensive.

So innate knowledge and a priori knowledge turn out to be co-extensive on their weakest definitions. This does not contradict Hart’s point that innate *beliefs* may be false. However, here too Hart has overstated his case. The innateness of a belief can count as evidence that it is a priori knowledge, and not just because innateness is a necessary condition on apriority. There are cases when it is reasonable to trust a belief precisely because it is innate. First, I think that in any situation in which evolutionary advantage would accrue if the process that yields innate belief were a reliable source of true beliefs it is reasonable to take this process to be a warranting one. Now all sorts of caveats must be placed on this assertion, as evolution is not the sort of engineer we often take it to be. We must remember that the organisms yielded by natural selection are not the ideal organisms, but merely the ones that are adequate to the environment. Further, we must

remember that even disadvantageous traits may be selected for if they are linked, either genetically or in the workings of the organism, to advantageous traits. All this means that there are plenty of reasons why our innate beliefs may be faulty, even when there is evolutionary pressure on them. However, on the whole, I think it is wise to trust instinct here. The other realm in which it is wise to trust instinct is the sphere of innate beliefs that will become true precisely because they are innate. If a universal grammar is innate, then the mere fact that it is innate will guarantee that all human languages will fit that grammar. Therefore the grammatical beliefs are guaranteed to be correct. Consider an infant with an innate universal grammar module. The infant will believe that her mother's yet undeciphered speech will conform to a certain pattern. But the very fact that this module is innate will guarantee that the mother's language will fit the pattern. Hence the innateness of the module guarantees its accuracy. Innate grammatical beliefs are self-fulfilling prophecies. Therefore, although Hart was right to say that innateness is not sufficient for apriority, he was quite wrong to think that it is irrelevant to it. There are cases in which it is reasonable to think that the mechanism of innateness is a warranting process.

So far we have only been talking about innate and a priori knowledge. However, both of these terms also apply to concepts, and it is worth seeing whether they are co-extensive in that arena as well. We said that *S* possesses a concept *C* a priori if in any possible world where she has the abilities typical of humans, and sufficient abilities to apply the concept *C* to an object, she would know whether the concept *C* applied to an object a reasonable percentage of the time. On the other hand, *S* possesses the concept *C* innately

iff it was true of her at birth that if she had the abilities necessary to manifest the possession of the concept *C* and if she developed in the way typical of humans, she would manifest possession of the concept *C*. The only difference between these two definitions is that the definition of innate concepts talks about manifesting the concept *C*, whereas the definition of a priori concepts talks about knowing whether the concept applies to an object. However, given that the most likely way one would manifest possession of a concept would be to apply it correctly to objects, it seems reasonable to say that these definitions are co-extensive.

Having outlined the dispositionalist nativist/negative apriorist position, the last thing we need to do is to consider the evidence for it. Since these definitions are weaker than the ones we previously considered, all of the old arguments will carry over. Indeed, they carry over in two ways. First, if they actually show the existence of representations and intuitions, then they have more than satisfied the requirements of the weaker definitions. But, second, it should be noted that some of the arguments given so far do not claim that there specifically must be intuitions or representations, they just argue for the general need for a priori or innate capacities. The argument for a priori mathematical knowledge based on the need for independent foundations for mathematics and physics fits this description.

The one remaining thing to do is to look at a second model of the empirical data on mathematical knowledge. The first model we looked at, the preverbal counting mechanism, fit well with the representationalist version of nativism. The second model we turn to now fits better with dispositional nativism. This model is based on the same

data—animal, infant, and childhood studies, reaction time and error rate studies, etc.—but assigns a weaker role to preverbal, analog representations. More importantly, there is a connectionist mock-up associated with this model that proposes that only a few aspects of these analog representations are present at birth. The remaining aspects of the analog system, however, develop spontaneously given normal input. For this reason, the model is not nativist under the representationalist definition of ‘innate’. It does, however, fit naturally with the dispositionalist definition.

Under the “triple-code” model, proposed by Stanislas Dehaene (1992), the brain uses three kinds of internal representation of number. The *auditory verbal code* uses general purpose language faculties to manipulate representations of number words in a natural language, e.g., ‘six hundred twenty two’ or ‘sex cent vingt-deux’, the *visual Arabic code* manipulates numbers in an Arabic format over a spatially extended representational medium, and the *analog magnitude code* operates using analog representations of the sort described by Gelman and Gallistel. The analog code is the only one that represents ordinal information—the only place where seven can be said to be more than six.⁴¹ Dehaene pictures this analog code as patterns of activation along an oriented number line. Like Gelman and Gallistel’s preverbal counting mechanism, Dehaene’s analog number code is shared by humans with other species, is already at work in the infant brain, and

⁴¹ Dehaene makes this point by saying that the preverbal code is the only code that carries “semantic information.” This phrase strikes me as claiming more than he wants to or needs to.

guides childhood acquisition of verbal arithmetic. Unlike Gelman and Gallistel's analog magnitudes, Dehaene's magnitudes are invariant. Distance and magnitude effects in processes that use the preverbal code are explained by saying that the internal number line is compressed logarithmically in the "greater than" direction. Each code has its own input and output mechanisms—reading and writing Arabic numerals for the visual code, hearing, speaking, and writing number words for the verbal code, and subitizing and estimating the size of large collections for the analog code. Transcoding mechanisms link all three codes internally. All the transcoding and input and output mechanisms are complex, involving both lexical and syntactic subsystems. Specific numerical operations can only take place in specific codes. For instance, comparison occurs in the analog code—this explains the distance and magnitude effects in number comparison. Recall of single-digit arithmetic facts, on the other hand, occurs in the verbal code, in keeping with the hypothesis that such facts are stored in rote, verbal memory.

Dehaene and Cohen (1995) have made a preliminary effort to localize the number codes and their associated mechanisms in the brain. Visual representations of number, including the parts of the verbal code involving written number words, are subserved by both hemispheres. The dominant mechanism is a sequence of maps belonging to the ventral visual pathway, which runs from the occipital lobe to the occipito-temporal region of both hemispheres, although the functioning of the right branch is severely limited. The analog code is also subserved by both hemispheres, using a mechanism in the vicinity of the occipito-parieto-temporal junction, including the inferior parietal region. The verbal code, on the other hand, only operates in the left hemisphere, using the areas commonly

associated with language. All the codes that operate in a given hemisphere are linked together, and the corresponding codes in each hemisphere are linked via the corpus callosum.

The existence of three separate codes with this pattern of localization is supposed to explain the deficit and split-brain data. The separation of the codes naturally predicts deficits that affect Arabic numbers but not ordinary words and written number words, or vice versa. The complexity of the input and output paths allows for selective deficits in either lexicon or syntax. The association of recollection of arithmetic facts with the verbal number frame and approximation and number ordering with the analog frame explains the numerous cases of preserved approximation and number reasoning with loss of recollection of arithmetic facts. The location of lesions is also generally consistent with the model. Deep number deficits come from insults to the inferior parietal, while less severe ones come from insults to linguistic areas and rote action areas like the basal ganglia. The research on split-brain patients seems to put most functioning in both hemispheres, save for verbal numbers and recall of arithmetic facts. Dehaene's model also works with his own PET scan experiment, which studied adults comparing and multiplying numbers mentally. Multiplication revealed activation in the inferior parietal region, as well as in the fusiform/lingual region and the left lenticular nucleus. The fusiform/lingual region was thought to be involved in transferring the visual input into the regions responsible for arithmetic fact retrieval or language production. The lenticular nucleus was supposed to be a part of the basal ganglular loop involved in rote fact retrieval. Several hypotheses were available for explaining the activation of the inferior

parietal, including the possibility that rote arithmetic fact retrieval was accompanied by an analog check. The only anomaly was the lack of significant activation of the inferior parietal region in number comparison tasks, which Dehaene attributed to the simplicity of the task.

In addition to giving preverbal arithmetic a weaker role in adult cognition, Dehaene has a different understanding of the way such arithmetic works. Dehaene, together with Jean-Pierre Changeux, have developed a connectionist mock-up of the mechanism they believe to be at work in animals and infants (1993). At the core of this model is a numerosity detector, which is able to identify the number of objects in a visual or auditory field, up to five objects, while discounting information about their size and location. The numerosity detector comprises three modules. The first is a 50-node one-dimensional input layer, or retina, where objects are represented as an area of activation. (The nodes are assumed to correspond to clusters of hundreds or thousands of actual neurons with common response properties, such as cortical columns.) The nodes of the input layer project onto a 9×50 array that normalizes for size and position. Each node in this array projects onto a line of 15 nodes of increasing thresholds of activation. The more activity in the normalization array, the farther activation will spread along this line. The final layer is another string of 15 nodes, connected to the previous string in such a way that activity along it will center in one of five areas, depending on the degree of activation along the previous line. Areas of activation along the final layer corresponding to different numbers overlapped, with closer numbers having greater overlap. The result was a distance effect when the system was set up to discriminate numerosities. The areas of

activation were also broader and flatter for higher numbers, which led to a magnitude effect. The discrimination task was accomplished by adding a two-node output layer, connected to the last layer of 15 nodes by connections of variable strength. The system was then trained to activate one output node when one numerosity was presented and the other when a different one was presented. The learning rule used was basically a Hebb rule with an additional factor that functions as a reward or punishment, stabilizing the system when it takes positive values, and destabilizing it when it takes negative values. Additional output nodes were added for more complicated tasks.

As it stands, Dehaene and Changeux's mechanism contains no ordinal information. According to Dehaene and Changeux, this is true to the state of a newborn's mind. A newborn essentially has unordered numerosity detectors. However, by the end of the first year of life, a person has developed ordinal knowledge. Dehaene and Changeux offer an extension of their model that is meant to explain how this transformation can occur. First they add to the model a short-term memory store, made up of highly autoexcitatory nodes, and a module that compares the short-term memory store with the current state of the numerosity detector. The next step is to add a system that can simulate the stimulation a child might receive playing with blocks and arranging them in piles. The play system consists of two "action" clusters, one of which is activated at random every 40 update cycles. Depending on which one is activated, an object is either added to or taken away from the input set (subject to the constraint that the input not exceed five objects). The network is then trained to guess which of the action nodes had been activated based on the current state of the numerosity detector and the state of its short-term memory. If the

network guesses correctly, the system is stabilized, if it guesses incorrectly, destabilized. As a result, the system is supposed to develop the knowledge that five is more than four, that three is more than two, etc. In the first experimental run, the system scored correctly 62.3% of the time, after a learning period of 300 cycles. In the second experimental run the setup was altered so that the action nodes could change the input by more than one object. The goal of the system was then to guess whether there had been an increase or a decrease. With this task, the system scored 74.8%.

It is possible to argue that the ability to discriminate number is representationally innate in this model, because the only training needed to manifest it modified the connection between the final layer of the numerosity detector and the output layer. Insofar as this ability is partially constitutive of possession of the concept of number, one could also say that the concept of number was partially representationally innate. But the full concept of number cannot be representationally innate because that would require ordinal knowledge to be innate. Ordinal knowledge will be dispositionally innate under this model, however. No doubt arranging blocks in piles is a part of the experience necessary to have a concept of number. But if our minds resemble Dehaene's connectionist model, we are born with a disposition to acquire ordinal knowledge given such experience. Therefore ordinal knowledge, and hence more of the concept of number, is dispositionally innate. Ordinal knowledge will, for the same reason, be a priori under a negative definition of the term.

There is thus an empirical case to be made for representationalist nativism, based on the preverbal counting mechanism originally proposed by Meck and Church. There is also

an a priori case for intuition-based apriorism, and an empirical case, based on the connectionist model of Dehaene and Changeux, for dispositional nativism and negative apriorism. The next chapter will respond to each of these cases separately.

Critique of Nativism and Apriorism

The previous chapter presented the case for mathematical nativism and apriorism. After examining various definitions of the terms ‘innate’ and ‘a priori’, the case turned into three separate cases. The job of this chapter is to reply to these three cases.

Critique of Representationalist Nativism

The representationalist nativist claims that some capacities—knowledge, concepts, or abilities—are innate because humans possess at birth a mechanism that represents those capacities. In the case of mathematical representational nativism, the mechanism was the preverbal counting mechanism. I will not argue here against the existence of the preverbal counting mechanism, although I have my doubts. Instead I will argue that the preverbal counting mechanism cannot provide pure mathematical knowledge, and if it cannot provide pure mathematical knowledge, it cannot provide pure mathematical concepts or abilities either. The argument will take the form of a dilemma: either the preverbal counting mechanism gives us capacities for applied mathematics, in which case it does not concern us, or it gives us the capacities of pure mathematics, in which case it cannot be wrong—and, as Wittgenstein teaches us, where we cannot talk about incorrect, we cannot talk about correct either.

To begin, consider a thought experiment. Imagine a preverbal counting mechanism that works differently than the actual mechanism. Suppose that every fifth time

a magnitude is supposed to pass into the accumulator one is taken out instead. This altered preverbal counting mechanism would therefore “count” using the sequence “1, 2, 3, 4, 3, 4, 5, 6, 5....”⁴² Two plus three would equal three, because the system would first count “1, 2,” then go on to count “3, 4, 3.” Similarly four and three would be five, because the system would count “1, 2, 3, 4,” then “3, 4, 5.” This hypothetical counting mechanism operates in all the same circumstances that the regular mechanism does. When infants see objects being grouped together, it causes the gate between the pacemaker to the accumulator to open once for each object. Similar things still happen when rats are trained to press a lever, and children are taught to count verbally.

The first thing that we should note is that even if all human preverbal counting mechanisms were altered in this way, normal arithmetic could still develop, both historically and in individuals. If humans retain the ability to put items in one-to-one correspondence and have the short-term memory and planning ability necessary to learn and execute simple algorithms, external objects can play the role that the preverbal counting mechanism normally does. Moreover, the abilities to put items in one-to-one correspondence and to learn and execute simple algorithms are part of the cognitive tool box that a human with an altered preverbal counting mechanism would have. These abilities would be a part of our hypothetical person’s cognitive toolbox simply because

⁴² One might think that this is impossible because the system would be unable to distinguish the first ‘4’ from the second. Therefore there would be no way for it to know when to go on to count ‘5’ and when to go back to ‘3’. However, we can assume that the accumulator can respond differently depending on whether four or six magnitudes have been fed into it, even this difference does not show up in the contents of the accumulator, and is hence unavailable for other cognitive mechanisms.

they are necessary for other skills typical of humans and animals generally. The ability to learn and execute simple algorithms is obviously necessary for any learning to take place whatsoever. The ability to put items in one-to-one correspondence is also a part of many basic skills. For instance Stanislas Dehaene (1997, 120) has pointed out that the notion of one-to-one correspondence is implicit in the idea of an exhaustive search. When an animal forages, it looks exactly once in each location food might be in an area.

If we grant people the ability to put items in one-to-one correspondence and to learn and execute simple algorithms, it is possible for them to develop normal arithmetic, although it will doubtless be more difficult. External objects can play the role that the preverbal counting mechanism usually plays. Indeed, this is more or less the way arithmetic actually developed. As I mentioned briefly in Chapter One, true arithmetic was probably developed in Mesopotamia by accountants who had to keep track of goods flowing in and out of the capital city of Uruk. This was first done using by moving piles of tokens representing quantities of goods from one bowl to another. Later, imprints of these tokens appeared on clay tablets. Then the imprint representing the quantity was separated from the representation of a particular kind of good. Number words, either pre-existing or freshly minted, were no doubt associated with these new number symbols. This system eventually evolved into proto-cuneiform. This path of the evolution of arithmetic is still open to humanity even if we alter their preverbal counting mechanism (or remove it, for that matter). All the work that is supposed to be done by the preverbal counting mechanism is done here by physical materials. Tokens act as magnitudes, bowls as accumulators. Physical number lines could have also severed the role Rochel Gelman

and C. R. Gallistel assign to a mental number line in learning arithmetic. Use of this system only requires putting collections of goods in one-to-one correspondence with collections of tokens, and to learn and perform simple manipulation of tokens. The use of these algorithms may have been historically guided by the preverbal counting mechanism, and conflicting intuitions from an altered preverbal counting mechanism may slow the development of arithmetic. However, a standardly functioning preverbal counting mechanism is not logically necessary for the development of arithmetic. This shouldn't be surprising, since the preverbal counting mechanism was originally introduced by analogy to the very physical mechanisms that are now substituting for them.

The evolution of a system of number independent of the preverbal counting mechanism shows that, even for a strict constructivist, a notion of number exists apart from the functioning of the preverbal counting mechanism. This allows us to talk about two ways in which the preverbal counting mechanism represents. I began Chapter 2 with the caveat that I was only talking about pure mathematical knowledge, not applied mathematical knowledge. I said that a rough way to distinguish the two was to stipulate that pure mathematical statements do not contain essential references to physical objects. This distinction can be applied to the preverbal counting mechanism. We can think of the preverbal counting mechanism as either embodying pure mathematical beliefs or applied mathematical beliefs. As an applied mathematical belief, particular states of the accumulator represent particular collections of physical objects in the world. Recall the definition of a representation we are working with: an organism represents an aspect of

the environment, “when there is a functioning isomorphism between an aspect of the environment and a brain process that adapts the animal’s behavior to it” (Gallistel 1990, 4).⁴³ It is perfectly natural to say that a particular state of the accumulator of the preverbal counting mechanism is isomorphic to a particular collection in the world and responsible for adapting the organism’s behavior in light of that collection. When we think of the preverbal counting mechanism this way, we are viewing the relevant isomorphism as holding between *tokens*. But we can also think of the preverbal counting mechanism as embodying a pure mathematical belief. Here, the preverbal counting mechanism in general represents some function that can be defined in verbal mathematics. Now the isomorphism is between types. The idea is that the way that the preverbal counting mechanism functions in general is isomorphic to some aspect of the environment, and adapts the organism’s behavior to that aspect of the environment. This target aspect of the environment can be minimally thought of as a way of functioning in the world, a way of manipulating objects exemplified by the accounting practices used at Uruk. Alternatively, it could be something like Kitcher’s “structural features of the world which allow us to segregate and recombine objects,” or even hypothetical Platonic objects. The distinction between pure and applied mathematics here is a distinction between representations of particular instances of mathematical operations and representations of general mathematical truths, not between different views of what those general mathematical truths could be.

⁴³ Recall also that the isomorphism might only be visible when the representation is interpreted as a part of a language, so that there needn’t be literal little pictures in our heads.

At this point we can frame a dilemma for the representationalist nativist. If we think that the preverbal counting mechanism embodies applied mathematical knowledge, it is capable of being in error. Indeed the modified version of it is in error. However, as applied mathematical knowledge, it is not a viable candidate for innate knowledge, nor is it the kind of mathematical knowledge we are concerned with. On the other hand, if the preverbal counting mechanism is supposed to embody a pure mathematical belief, it is incapable of error, and as a result cannot be knowledge. Now the first horn of this dilemma is fairly easy to understand. When we view the preverbal counting mechanism as embodying applied knowledge, we are thinking of individual states of the accumulator as representations of particular collections in the world. If this is the case, then the altered preverbal counting mechanism clearly misrepresents the world. If a person perceives five objects being placed into a box, the preverbal counting mechanism will only represent the box as holding three objects. However, this kind of applied mathematical belief is not a viable candidate for innate knowledge. It would certainly be strange if knowledge about a particular collection of objects was present in us at birth. Fortunately, interpreting the preverbal counting mechanism as embodying beliefs about particular collections entails no such thing. The representation of a particular collection is not formed in the preverbal counting mechanism until the individual comes into contact with the particular collection. So although error is possible if the preverbal counting mechanism embodies beliefs about particular collections, innate knowledge is not possible. Moreover, even if it were possible, it would not be the kind of knowledge we are interested in. As I noted at the

start of Chapter 2 my concern is only with pure mathematical statements, as these are the only statements that can reasonably be said to be normative and not descriptive.

The second horn of the dilemma is a little trickier. As a pure mathematical belief, the altered preverbal counting mechanism must be trying to represent some operation, but what operation? It could be an attempt to represent ordinary counting, in which case it is in error. On the other hand, it could be an attempt to represent the altered form of counting where one periodically subtracts one instead of adding one, in which case it is entirely successful. But if the altered preverbal counting mechanism is attempting to represent altered addition, then there is no way that any kind of preverbal counting mechanism can ever be in error. No matter how it is altered, there will always be a function that it successfully models. The situation is a lot like the impossibility of universal error Wittgenstein describes in book 1 of the *Remarks on the Foundations of Mathematics*, §§135–37. There he noted that if everyone consistently got a different answer for a calculation than they do now, if for instance everyone thought 12×12 was 145, we would simply declare 145 to be the new correct answer to the problem. We would adopt $12 \times 12 = 145$ as a part of our new definition of “ \times ” and use some other symbol for the function that maps 12 and 12 onto 144. Similarly, as long as we are altering the preverbal counting mechanism so it behaves consistently, there will always be a function that we can say it is successfully representing. So the question is whether we should view the altered preverbal counting mechanism as representing the altered function correctly, in which case error is impossible, or the standard function incorrectly.

The kind of counting that the altered preverbal counting mechanism is trying to represent depends on how you think representational content is determined. On old-fashioned Lockean views of representation, the preverbal counting mechanism will represent whatever function it is isomorphic with. On the Lockean view, things represent by resembling, and an altered preverbal counting mechanism will most closely resemble an altered function. What the representationalist needs in order to create the possibility of misrepresentation is a way to tie the content of a representation to the aspect of the world it is responding to. Gallistel's definition of representation seems to offer us such a link. In Gallistel's definition, the content of a representation is determined by two factors. First there must be an isomorphism between the representation and an aspect of the environment, and second the representation must adapt the organism's behavior to that aspect of the environment. It is easy enough to say from here that the representation is trying to represent the part of the environment that it is adapting the organism's behavior to, and it is successful if it is isomorphic to that part of the environment.⁴⁴

But how do we know what aspect of the environment the representation is adapting the organism's behavior to? Specifically, how can the representationalist nativist rule out

⁴⁴ An alternative might be to say that the isomorphism fixes the referent and the adaptation determines veracity. Presumably, we would say that a representation is veridical if it adapts our behavior well to an aspect of the world, where 'well' is defined in Darwinian terms, à la Ruth Garrett Millikan. However, this will not help us say that the preverbal counting mechanism can be in error if it is taken to embody pure mathematical beliefs. If the isomorphism fixes the referent, then however the preverbal counting mechanism is altered, it refers to that altered function. Presumably, it also prepares the organism well to deal with situations that obey that altered function. The fact that it takes elements of the environment to be instances of the altered function, when in fact they are only instances of ordinary accumulation, is not relevant.

the possibility that the altered preverbal counting mechanism isn't supposed to adapt the organism's behavior to situations where objects are being placed into an accumulator, and then occasionally removed from it? If this could be construed as the target situation, preverbal counting mechanisms would again be incapable of error. The natural thing to say is that the representation is adapting the organism's behavior to an aspect of the environment if the representation is caused by the aspect of the environment and is used to modify the organism's behavior. Whether a representation is caused by an aspect of the environment would be judged by whether the representation occurs if and only if the aspect of the environment is present. This would work fine if we thought the preverbal counting mechanism embodied knowledge about particular collections. If the preverbal counting mechanism embodies applied knowledge, each state of the accumulator is an individual representation. That state is caused by the presence of an individual collection, which would then be the aspect of the environment the state of the accumulator is representing. The state of the accumulator is representing the collection accurately if it is isomorphic to it. The suggestion breaks down, however, when we think of the preverbal counting mechanism as embodying general mathematical knowledge. The preverbal counting mechanism will not just be activated in situations where objects are accumulating normally. It will also count in situations where objects seem to be accumulating because of bad light, great distance, etc. Which of these situations is the kind of situation the organism is adapting itself to? The same function could be a good attempt to represent some of them and a bad attempt to represent others. It is tempting to say here that the organism is attempting to adapt itself to all of them, and only sometimes

succeeding. But to say this is to revert back to saying that the preverbal counting mechanism represents individual collections, rather than embodying pure mathematical knowledge. As soon as one wants the preverbal counting mechanism to represent a general mathematical truth, one has to find a general feature of the situations it responds to for it to be isomorphic with, and no such feature is present in all the situations.

The situation is similar to the disjunction problem with causal accounts of representational content discussed by Jerry Fodor (1987), and solutions from that discussion might be applicable here. The causal theory of representation, in its crudest form, says that a mental structure represents an aspect of the world if it is caused by it. Crude causal theory is clearly inadequate. Demanding a reliable causal link between horses and ‘horse’ tokenings means that all horses and only horses must lead to tokenings of the representation ‘horse’. But, as Fodor points out, there are problems with both the ‘all’ part and the ‘only’ part. Not all horses cause ‘horse’ tokenings—some go unseen. Moreover not all tokenings of ‘horse’ are caused by horses—in bad light or at a distance, one might mistake a cow for a horse. ‘The disjunction problem’ is another name for the problem with saying that *all* horses cause horse tokenings. Under the crude causal theory, the token ‘horse’ seems to refer to an extended disjunction, “horse or cow or full scale horse replica or ...” Moreover, it seems to always refer veridically, since whenever it is tokened, it is tokened by a part of the disjunction. Fodor’s solution is to appeal to the idea that falsity is dependent on truth in a way that truth is not dependent on falsity. According to Fodor, we can say that a statement is false only by contrast with a true statement, but true statements bear their truth independently of false ones. As Baruch Spinoza put it,

“Even as light displays both itself and darkness, so is truth a standard both of itself and falsity” (*Ethics*, scholium to bk. II, prop. 43). Applied to representations, this general principle implies that the situations where cows lead to tokenings of ‘horse’ are dependent on the situations where horses lead to tokenings of ‘horse’, in a way that situations where horses lead to tokenings of ‘horse’ are not dependent on situations where cows lead to tokenings of ‘horse’. Think of it this way: if I see a cow, and that causes me to think ‘horse’, It’s only because I have been trained to think ‘horse’ when I see horses. On the other hand, when horses cause me to think ‘horse’, this has nothing to do with cases where I might mistake a cow for a horse. Fodor calls this *asymmetric dependence*.

Gallistel’s definition of representation does not imply the causal theory of representation Fodor describes, but as we have interpreted Gallistel, his definition does have a causal element. Gallistel’s representations have two elements, an isomorphism between a brain structure and an aspect of the world, and a mechanism that allows the organism to adapt its behavior to that aspect of the world on the basis of that representation. We fleshed this out by saying that the adaptive mechanism fixes the referent, and the isomorphism determines veracity, and further that the adaptive mechanism was a kind of causal link between referent and representation. A situation similar to the disjunction problem came up when we tried to say that the preverbal counting mechanism in general (rather than particular states of the preverbal counting mechanism) represented a pure mathematical function (rather than a particular collection in the world). The preverbal counting mechanism responds to a variety of situations, not all of which are instances of proper accumulation, just as the token ‘horse’ might be

occasioned by either horses or cows. So what function is the preverbal counting mechanism trying to represent? If we were to adopt the asymmetric dependency solution, we would say that the response of the preverbal counting mechanism to the inappropriate functions is dependent on its response to the appropriate ones, and that the converse dependence does not hold. But how are we going to find this asymmetry? Note that we cannot appeal to the idea that, in general, falsity is dependent on truth but not vice versa. Truth and falsity are not at issue here. What is at issue is what aspect of the environment the organism is adapting its behavior to. One might try an evolutionary gambit, and say that the organism evolved to respond to situations involving ordinary accumulation of objects, and would not be responding to the spurious cases if it had not evolved to respond to the proper ones. But this will not necessarily place the line between correct and incorrect where we want it. Often the mechanisms that organisms use to filter perception—say, to highlight signs of predators—are selected to be oversensitive because the costs of a false positive are significantly lower than the costs of a false negative. Similarly, there may have been an advantage to the preverbal counting mechanism responding to situations that are not proper counting. So again we are left with no way to identify what pure mathematical function the preverbal counting mechanism is trying to represent, unless we simply say that it represents the function it is isomorphic to, in which case it cannot be wrong.

I have been assuming that if the preverbal counting mechanism cannot be wrong, it does not give knowledge. My basis for this is the Wittgensteinian dictum that where we can't talk about incorrectness, we cannot talk about correctness either. More needs to be

said about this idea, if only because for centuries philosophers took the opposite idea for granted: only knowledge gained by infallible means was real knowledge. But the sort of infallibility desired by foundationalist epistemology is not the sort we have here. Foundational beliefs were typically beliefs humans had necessarily and which necessarily matched the way the world was. With the preverbal counting mechanism, beliefs can be conceivably changed in any number of ways. The beliefs remain true only because we can construct a new function for them to be true of. This indicates that, as a pure mathematical belief, the preverbal counting mechanism is contentless. If any aspect of it can be changed, then none is important.

If the preverbal counting mechanism cannot bring pure mathematical knowledge, it cannot bring pure mathematical concepts or abilities either. So far, we have been describing the preverbal counting mechanism as embodying *knowledge*, either knowledge of the state of a particular collection or knowledge of a certain function. But we have also said that the preverbal counting mechanism embodies a mathematical concept, the concept of number, and an ability, the ability to count. These other capacities exhibit the same pure/applied split that mathematical knowledge did. The preverbal counting mechanism embodies a pure concept of number in that its functioning in general is isomorphic to the structure of number in general. But we can also see particular states of the preverbal counting mechanism as embodying a concept about a particular collection, a numerical haecceity, if you will. Similarly, the know-how the preverbal counting mechanism embodies could be knowledge of how to count in general, or knowledge of how to count a particular collection. Since these mathematical capacities exhibit the same

pure/applied duality that mathematical knowledge did, we can frame the same dilemma about them. And if one accepts the argument given for mathematical knowledge, the result is the same. Either the concepts and abilities in question are of no concern to us, or they are incapable of error.

Critique of Intuitionist Apriorism

The intuitionist apriorist defined a priori knowledge as knowledge that came through a specific channel, the intuition. Mathematical knowledge is a priori because it is based in part on knowledge gained through a mathematical intuition. In Chapter 3, I argued that this mathematical intuition plays two roles: It secures the validity of inferences and establishes the truth of axioms. I also argued that the most plausible form for this intuition was the hybrid Kantian/Husserlian form described by Charles Parsons (1979–80/1996). Under this model, intuitions of types are founded on perceptions of tokens, à la Husserl, but the intuitions are insights in to the necessary structure of our perception, à la Kant. I ended up endorsing three arguments for the existence of such an intuition. The first two were from Larry BonJour, and both involved an appeal to the need for metajustification. The first argument claimed that we needed to be able to provide a metajustification for ordinary inferences in a way that could not be written as a new premise. Only an intuition, it was thought, could do this job. The second argument claimed that the alternatives to apriorism, the moderate empiricism of the positivists and the radical empiricism of Quine, were incoherent. The problem with the positivists was the familiar charge that they cannot produce an adequate definition of analyticity. It was

the problem with Quine that brought in the need for a metajustification. Quine, according to BonJour, needs to provide a metajustification for his epistemological principles, but cannot. The final argument for an intuition from the last chapter that I accepted was the argument based on need for mathematics to have a foundation apart from the empirical sciences. My basic strategy in this section will be to deflect each of these arguments, showing how they actually support a Wittgensteinian conclusion.

Both of BonJour's arguments rely on a demand for a metajustification. But the demand for a metajustification is self-defeating and ultimately forces Wittgensteinian conclusions on us. This is easiest to see in the inference argument. In the inference argument, BonJour envisions a situation essentially like the situation Wittgenstein imagined in the regress of interpretation argument, or Lewis Carroll described in "What the Tortoise said to Achilles" (1895). The parallel to Carroll is the most straightforward. The tortoise, just like BonJour, asks us to imagine an arbitrary inference and to explain why it is justified. He then points out that this justification can be written as a new premise. The inference as a whole then requires a new justification. In Carroll's version, the subsequent regress is simply presented as a paradox. BonJour wants us to think that the only way to stop the regress is to invoke an a priori intuition. When Wittgenstein turns to this kind of situation, however, he tries to prevent the regress from launching in the first place. As we saw in Chapter 2, Wittgenstein's regress of interpretation argument was a part of a larger effort to show that nothing can stand behind the applications of a rule, deciding which are correct and which incorrect. The correctness and incorrectness must be an irreducible feature of the actions themselves. The regress of interpretation

manifested this idea by asking us to imagine that a mental image guides our applications of the word 'cube'. Supposedly we would draw mental lines of projection from an image of a cube to the actual cube. But, Wittgenstein points out, there are alternative ways of drawing lines of projection, and the only thing that could rule them out would be another mental image and scheme of progression. Because this regress is unstoppable once it starts, Wittgenstein must stop it before it starts, and the way he does this is to declare 'correct' and 'incorrect' are irreducible features of the actions themselves.

In essence, we are being asked to choose between two ways of stopping the regress of interpretation. The contrast between these methods is sharp. Wittgenstein describes the conclusion of his rule-following arguments this way: "There is a way of grasping a rule that is *not* an *interpretation*, but is exhibited in what we call 'obeying a rule' and 'going against it' in actual cases" (*PI* §201). This idea of obeying a rule in practice is surrounded by images that emphasize how automatic the process is. When Wittgenstein's interlocutor asks him how he knows how to follow a rule for extending a series of numbers, he replies "If that means 'have I reasons?' the answer is: my reasons will soon give out. And then I shall act, without reason" (*PI* § 211). In the next passage, he adds this illustration, "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" (*ibid.*, §212). Additional remarks reinforce these themes: "If I have exhausted the justifications I have reached bedrock, and my spade is turned. I am inclined to say, 'this is simply what I do'" (*ibid.*, §217); "When I obey a rule, I do not chose/ I obey the rule blindly" (*ibid.*, §219). The thrust of these

remarks is clear. The correctness of an action cannot be rationalized; it is best seen in cases where one acts without thinking.

The intuition that the apriorist want to stop the regress is, by contrast, a specifically intellectual act. Under Parsons' hybrid Kantian/Husserlian model, formal intuitions are acts of consciousness founded on acts of perception. When Husserl describes a founded intuition he is quite explicit in describing it as a separate act of consciousness, laid on top of ordinary perception. Husserl's idea was that we grasp ideal entities by grasping instances of them. But this formal grasping is very different than the comprehension that occurs in perception. Perceptual intuition may be very complex. A perceptual object will appear as a sequence of perspectives over time, for instance. But all of these perspectives are united under a single act of consciousness: "In this unity, our manifold acts are not merely fused into a phenomenological whole, but into *one act*, more precisely into *once concept*" (1900–1/1970, §47). These perspectives fall under one act because they are perspectives on one object. By contrast, the formal intuition creates new entities, apart from the objects of perception, and as a result require separate acts. "In the sense of narrower 'sensuous' perception, an object is directly apprehended or is itself present...It is not *constituted* in relation, connective or otherwise articulated acts, *acts founded on other acts which bring other objects to perception*" (ibid. §46). In categorical intuition, "what we have are acts which, as we said, *set up new objects*, acts in which something *appears as actual and self given*, which was not given and could not have been, as what it now appears to be, in these foundational acts alone" (ibid.). Husserl's idea is clear, the categorical intuition is an intellectual, synthesizing act, separate from perception.

The separation of the categorical and perceptual intuitions prevents a founded a priori intuition from stopping the regress of interpretation. Because there is a gap here, there are multiple ways a categorical intuition may be founded in a perceptual one, multiple ways of synthesizing the same percept. Because we have opened the door to requests for metajustification, one of these multiple ways of synthesizing must be justified over the others, and the regress begins again. Consider one of the examples Parsons accepts as a legitimate case of a founded intuition, the stroke language developed by David Hilbert. This language has one symbol, '|', and well-formed expressions are strings that contain only this symbol. Two tokens of this language belong to the same type if they can be placed side by side so that the strokes correspond one-to-one. Parsons felt this shows types could be intuited as part of ordinary perception, and that this intuition of types could be easily turned into an intuition that certain propositions are true, for instance that ||| is the successor of ||, or that the stroke series can be extended indefinitely. Now it is not hard to generate Wittgensteinian situations here. Someone could say that ||| and |||| belong to the same type because they have a deviant notion of one-to-one correspondence. The Husserlian background to Parsons' intuition even gives us a concrete way of describing what the deviant is doing here: She is founding a different categorical intuition on the same perception. Similar examples can be developed for the claim that ||| is the successor of || or that the sequence can be extended indefinitely. The question we can now put to the apriorist is, "Why is your categorical intuition justified and the deviant one not?" The apriorist might appeal to the nature of one-to-one correspondence, perhaps brandishing a diagram illustrating what one-to-one correspondence is supposed to look like. But this

will only start the regress of justification going again, and this was exactly what the a priori intuition was supposed to stop.

On the other hand, Wittgenstein's notion of acting without justification, but not without right, avoids all these problems. There is no separation between the correct action and its correctness that would allow for multiple interpretations. This kind of correct action can be thought of as a kind of know-how. Calling it know-how does not open up any new options for the mathematical apriorist. As I pointed out in the last chapter, Ryle has argued successfully that know-how cannot be a priori. In fact, the same aspect of know-how that keeps it from being a priori allows it to stop the regress of justification: there is no justification of know-how. For there to be a justification of know-how there would have to be a difference between knowing-how and believing-how. But we do not speak of believing-how. This prevents know-how from being a priori, since apriority is a matter of justification. It also is what allows us to stop the regress of justification, since the regress is formed by the distance between justification and action.

The apriorist's second argument for the existence of an intuition also depended on the demand for metajustification. In this case, a request for a metajustification was laid at Quine's feet. In fact, additional moves made by the apriorist in the argument against Quine might be applied back to the inference argument, to strengthen it. These moves will not strengthen the inference argument enough, however. The demand for a metajustification from Quine and the strengthened inference argument fall victim to the same replies that the first version of the inference argument fell victim to. The demand for metajustification backfires, because once the demand was made, it could not be

stopped. This led to a Wittgensteinian conclusion, because it was seen that the regress had to be ruled out before it started.

Part of the argument for apriorism was an attack on various thinkers who did not believe in an a priori intuition, including Quine. The claim against Quine, in particular, was that he had numerous general epistemological principles—conservatism, the belief that mathematical and logical statements were in the center of the web of belief—that he could not justify. The apriorist, by contrast, could offer a metajustification for her epistemic principles. The metajustification took a Kantian form: if we were not justified in relying on our a priori intuitions, experience would be impossible. The argument depended on the existence of an intuition, but that is all right in this context. In the face-off against Quine, we are presented with two epistemological world views, one of which is self-validating, and the other of which isn't. This Kantian metajustification might also be applied back to the inference argument. The inference argument claimed that an a priori intuition was necessary to justify ordinary inference. The Wittgensteinian reply was that an intuition could not rule out the demand for another metajustification, because it could not specify the way it was to be applied. The Kantian riposte is now that the intuition must be able to specify its own application, otherwise experience would be impossible. This reply is weak, because it seems to only strengthen the need for a halt to the regress, rather than actually providing the intuitionist a way to halt it. This is just a symptom of a deeper problem, however, namely that the Kantian metajustification simply doesn't work in the face of Wittgensteinian considerations.

If the Kantian metajustification went through, it would effectively stop a regress of justification from forming. Depending on how you look at it, this regress will be stopped one of two ways. Either the halt is complete because no further questions can be asked, or further questions can be asked, but they all have the same answer, so the demand for metajustification results in a “turtles all the way down” situation. One might ask, after being told that the deliverances of the intuition must be reliable in order for experience to be possible, why we should rely epistemically on principles that are necessary for experience to be possible. Some might not consider this a legitimate question, in which case the regress of justification stops right here. Others might take it to be an odd question, but nevertheless one that must be answered. If you take this angle, then the regress of justification becomes a “turtles all the way down” regress. The answer to the question “why we should rely epistemically on principles that are necessary for experience to be possible?” is simply “because if we don’t experience will be impossible.” This will also be the answer for any further justification demanded. The world rests on the back of the turtle, but beneath that, it is turtles all the way down. Either way you look at it, the Kantian metajustification of the deliverances of intuition will halt the regress of justification if it goes through.

Unfortunately for the apriorist, the kind of Wittgensteinian situations we were just looking at show that the intuition offered by the apriorist is not necessary for experience to be possible and cannot explain how experience is possible. The intuition presented by the apriorist cannot explain how experience is possible because it cannot stop the regress of justification. This doesn’t just mean that the intuition is not *sufficient* to explain how

experience is possible. It cannot even make a contribution to that explanation. When we ask the question “how is experience possible?” we are asking how we can perceive an ordered world, a world that it is possible to make some sense of. An intuition cannot help to answer this question because it contributes nothing to our understanding of the ordering of events. But instead of telling us whether something fits or doesn’t fit in a pattern—whether two tokens are of the same type, or whether $|||$ is the successor of $||$ —it hangs alongside these judgements, just as Wittgenstein described an interpretation as hanging alongside correct action. The intuition is also not necessary for experience to be possible. To understand how experience is possible, one needs to understand instead how an action can be right without being justified. Once we realize that a judgement does not always need justification, explaining the order of experience is no trouble. We can explain order up to a point, and then just stop when justification runs out.

The final argument for apriorism we need to look at is the need for an independent foundation for mathematics. The idea was that if mathematics does not have its own foundation of intuition-based knowledge, it would become merely a very abstract branch of empirical science. This would, of course, be perfectly fine for people like Quine and Kitcher, but others find it counterintuitive. “Isn’t it odd,” we saw Penelope Maddy ask in the last chapter, “to think of ‘ $2 + 2 = 4$ ’ or ‘the union of the set of even numbers with the set of odd numbers is the set of all numbers’ as highly theoretical principles?” (1990, 31). Arguments like this are appealing in a climate dominated by Quine’s empiricist views of mathematics. In Michael Resnick’s (1997) terminology, one is either a holist, lumping mathematics and empirical science together like Quine does, or a separatist, in which

case one will be led to believe in mathematical intuitions. The picture changes dramatically when the normative approach to mathematics is introduced. In a broad sense, Wittgenstein's view is clearly separatist, since it grants mathematical statements a fundamentally different status than empirical ones. However it also presents a unique view of how mathematics and empirical science interact, the latter structuring the former. In any case, the arguments against holism put forward by Maddy and others cannot be said to point unequivocally to a mathematical intuition. They can just as easily point to a normative view of mathematics.

Critique of Dispositionist Nativism and Negative Apriorism

The combined dispositionist nativist and negative apriorist view holds that mathematical knowledge and concepts are innate and a priori in a very weak way. Mathematical knowledge is innate and a priori in that at birth, a certain complex counterfactual is true of us: in any possible world where we have the capacities typical of human beings, and enough experience to understand the concepts involved in these mathematical propositions, we would believe them, and have a warrant or justification for believing in them. Mathematical concepts are innate in that a slightly different counterfactual is true of us at birth: in any possible world where we had the abilities typical of humans and the abilities necessary to manifest possession of the concept in question, we would be able to tell whether the concept applied to a reasonable proportion of given objects. The evidence that mathematical knowledge and concepts are innate in this sense is simply the evidence, both empirical and rational, that was produced by the

other versions of nativism and apriorism, along with an additional, connectionist empirical model, which is better suited to the dispositionalist approach to nativism than to the representationalist approach.

Because the hybrid nativist/apriorist view borrows arguments from the previous two positions, some of the critiques of those positions will also carry over. All of the arguments against intuitionist apriorism clearly carry over. The argument against representational nativism challenged the claim that a specific representational mechanism could bring knowledge, concepts, or abilities. To the extent that the hybrid nativist/apriorist relies on this model, the argument from that earlier section also carries over. However, the argument in the earlier section did not challenge any of the empirical data presented by the nativist. Since these data also support other models of mathematical knowledge, including the connectionist model well suited to dispositionalism, they continue to provide support for the hybrid nativist/apriorist view. So as things stand now the argument for dispositional nativist and negative apriorism rests on the empirical evidence insofar as it supports the connectionist model.

The hybrid position is sufficiently weak to be completely compatible with Wittgenstein's claim that mathematical statements are expressions of norms. At first this may seem absurd. The position I have outlined is based largely on empirical research. It attempts to explain the origin of knowledge by looking at the development of the brain. But didn't Wittgenstein claim repeatedly that you cannot explain mental phenomena by appealing to brain states? In general, isn't Wittgenstein the enemy of empirical psychology? To see why Wittgenstein would not object to the dispositionalist

nativist/negative apriorist view, I will first of all look closely at the things Wittgenstein actually said about the mind and the brain. Then I will argue for the compatibility of both Wittgenstein's view of the brain and Wittgenstein's view of mathematics with the hybrid nativist/apriorist view.

The two main sources for Wittgenstein's beliefs about the mind and the brain are the *Philosophical Investigations* and *Zettel*. In both books he argues that one cannot explain mental states in terms of brain states, and in each case, the core of the argument is that such an explanation would reduce normativity to empirical phenomena. In the *Philosophical Investigations*, Wittgenstein makes the point three times, once in the context of knowledge, once in the context of reading, and once in the context of speaking silently to oneself. The claim gets a more sustained treatment in *Zettel*, the published version of a collection of remarks Wittgenstein clipped from his notebooks, arranged in bundles and kept in a box for future use. A stretch of remarks in that work, §§605–13, elaborate on the claims about the brain made sporadically in the *Investigations*, and present a clearer argument for them. The notebooks that are the sources for these remarks have also been published, however they contain little that is not in *Zettel*, so our discussion will focus on *Zettel* and the *Philosophical Investigations*.

Both the discussions of mind and brain in the *Philosophical Investigations* and *Zettel* take place against the background of Wittgenstein's belief that knowledge was a kind of

disposition.⁴⁵ Wittgenstein's motivation for thinking this was the way knowledge interacted with time. The treatment of the issue in the *Philosophical Investigations* begins when he breaks off a discussion of how to apply a rule to ask, "When do you know that application? Always? Day and night? Or only when you are actually thinking of the rule?" Concerns about the temporal aspect of attributions of knowledge recur in the subsequent passages. In §151 he points out that there is one instance where we say that knowledge is temporal: when we suddenly understand something, we say, "Now I know!" Wittgenstein also inserted two remarks written on slips of paper into his typescript at this point. One asked the reader to compare the grammar of, "Since yesterday I have known this word," and sentences like "He was depressed the whole day." The other insertion asked if one only knew the rules of chess while one was making the moves, and quipped, "How queer that knowing how to play chess should take such a short time, and a game so much longer!"

The discussion of knowledge and dispositions in *Zettel* also focuses around time. The segment begins with a general remark about how removed abilities are from actual events: "Being able to do something seems like the shadow of actual doing, just as the sense of a sentence seems like the shadow of a fact, or the understanding of an order the shadow of its execution. In the order the fact as it were 'casts its shadow before'. But this shadow, whatever it may be, is not the event" (Z §70). The subsequent remarks make it

⁴⁵ This belief alone, of course, does not make him a dispositionalist nativist. As we said in the last chapter, if innate knowledge is a kind of disposition, it is actually a disposition to have a disposition.

clear that the shadowy nature of understanding is a product of its relationship to time. *Zettel* §71 is essentially the same as the remark on the first scrap of paper Wittgenstein slipped into the discussion of knowledge in the *Philosophical Investigations*, where he asked us to compare the grammar of “Since yesterday I have understood this word,” and expressions like, “He was depressed the whole day.” Remarks after *Zettel* §71 look at the way attention interacts with knowledge and time. Because a sensation persists in time, it is possible to attend to its duration. On the other hand, one cannot attend to the duration of knowledge (*Z* §75–77, 81–2).

The best response to the oddities of the relationship between knowledge and time is to view knowledge as a kind of disposition. Dispositions, like pieces of knowledge, have definite points of inception. One can say, for instance, that a person’s inclination to treat animals with compassion dates from watching a sentimental documentary on the feelings of animals. On the other hand, dispositions have a rather vague existence after their inception. Although we can say for sure that a disposition to a kind of behavior exists when the behavior is manifest, it is hard to say when else it exists. In keeping with the practice established in Chapter 2, I would like to call the thesis that knowledge is a kind of disposition an *explanation* of the strange relationship attributions of knowledge have to time. This is, of course, not at all how Wittgenstein views the claim. He only uses the term “disposition” in the *Philosophical Investigations*. There he presents the idea as if it were a clarification of what we are ordinarily inclined to think. After he first presents the problem of time and knowledge, he suggests that people commonly view knowledge as a state of mind, but “if one says that knowing the ABC is a state of the mind, one is

thinking of the 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” (*PI* §149). Wittgenstein never challenges the idea that knowledge is a kind of disposition, and his later remarks hint that he accepts this view. For instance in §151 he says, “The grammar of the word ‘knows’ is evidently closely related to that of ‘can’, and ‘is able to’. But also closely related to that of ‘understands’. (‘Mastery’ of a technique.)” Abilities are much more clearly like dispositions than knowledge, in that they can be either dormant or active. Although Wittgenstein’s relationship to this thesis is tentative, for my reconstruction of his argument I will go ahead and assert it on the grounds that it provides the best explanation of the relationship between knowledge and time.

The thesis that knowledge is a disposition is the chief premise of Wittgenstein’s first argument that mental states cannot be explained in terms of brain states. A disposition, Wittgenstein claims, cannot be reduced to a brain state. This claim is introduced in 149, almost at the same time as he introduces the idea that knowledge is a disposition. As we noted, he begins that passage by saying that people are inclined to think of knowledge as, “the state of a mental apparatus (perhaps the brain.)” But if knowledge is a disposition “there are objections to speaking of a state of mind here, inasmuch as there ought to be two different criteria for such a state: Knowledge of the construction of the apparatus, quite apart from what it does” (*ibid.*). One can determine the state of a machine by examining the configuration of its parts, but to determine what a machine does, one must see it in action. There are therefore different criteria for determining the function of an entity and its state. But a disposition, Wittgenstein is telling us, is a functional entity. So

if, when we call a disposition a state of mind, we are thinking of something like a configuration of the brain, we are simply making a category mistake. A similar argument could be made for other mental states if they are viewed as being like dispositions. If the mental phenomenon can only be identified through the activity of the organism, then they will be categorically different than aspects of the physical construction of the organism. For instance, being angry can be seen as a disposition to make certain kinds of statements. If anger can only be identified through behavior, then it is categorically different from any kind of brain state, which could be identified by looking at the organism statically.

This first effort to distinguish mental states and brain states is incomplete, at best. For simple machines, at least, we can identify dispositions by examining the configuration of the parts. If the snare on a mousetrap is pulled back so there is tension on the spring, the trap will be disposed to snap shut if the bait is disturbed. Why can't there be physical states similarly related to the dispositions we call knowledge? Similarly, why can't we find physical states that correspond to any mental state? Wittgenstein's answer is that we might very well be able to find a physical state that allows for a disposition or mental state. However, being in such a physical state could never serve as a *criterion* for having that disposition or mental state. For Wittgenstein, 'criterion' is something of a term of art. In the *Blue Book*, Wittgenstein contrasts criteria with symptoms, pointing out that while an inflamed throat might be a symptom of angina, the presence of a certain bacterium in his blood actually specifies the nature of the disease (*BB*, pp. 24–25). Wittgenstein elaborates, "To say 'a man has angina if this bacillus is found in him' is a tautology or it

is a loose way of stating the definition of 'angina'. But to say, "a man has angina whenever he has an inflamed throat' is to make a hypothesis" (ibid., p. 25). Based on this, it is safe to say that a criterion must be a part of the definition of the thing it is a criterion for. Therefore, although it may not be fair to say that 'criterion' simply means 'necessary condition', it is safe to say that at least part of being a criterion is being a necessary condition. So Wittgenstein claims that no physical property could be necessary for a disposition or mental state. The physical state we identify will never be the only physical state that corresponds to the mental state, nor will we ever be able to draw up a list that we know to be complete of all the physical states that correspond to the mental property. Therefore, a physical state can never be a criterion for a mental state. But Wittgenstein's original point was that brain states and dispositions must be different kinds of things because they have different criteria. Therefore his original point stands.

This argument depends, of course, on Wittgenstein showing that there can be no necessary connection between a physical state and a disposition or other mental state. Wittgenstein's second statement in the *Philosophical Investigations* of the thesis that mental phenomena cannot be explained by physical structures doesn't offer a reason for thinking that there cannot be a necessary connection between mental phenomena and physical structures, it merely hints that there cannot be one. While considering the possibility that a physical criterion could be used to determine when someone is reading, Wittgenstein wonders if the connection between the physical criterion and reading is "a priori" or "only probable" (*PI* §158). The reader of the *Philosophical Investigations* is only given reasons to think that there can be no necessary connection between a brain

state and a mental state in §376. There Wittgenstein asks what criterion could be used to determine that he and someone else were both saying the same sentence silently to themselves. “It might be found that the same thing took place in my larynx and in his,” Wittgenstein writes, “... But when did we learn the use of the words ‘to say such-and-such to oneself’ by pointing to a process in the larynx or the brain? Is it not also perfectly possible that my image of the sound *a* and his correspond to different physiological processes?” (*PI* §376). The criteria for the correct use of a term or phrase are always given somehow during the process of learning to use that term or phrase. If they weren’t, we would never learn language. But in teaching the phrase ‘to say such-and-such to oneself’ we do not use people’s brains in certain configurations as examples. Something similar can be said about the terms “knowledge” and “reading.” In each case we do not appeal to the shape of the brain in teaching the word—we don’t even know what brain states to use! This means that any connection we make between having a certain brain state and knowing, or reading, or speaking silently to oneself, will be accidental. We can make such connections by applying the criteria for knowing or reading or speaking silently to oneself and seeing that they end up demarcating the same phenomena as the criteria used to judge whether a person is in a specific brain state. However, we cannot merge the two sets of criteria. We could theoretically begin to teach children the meaning of knowledge by pointing to MRI scans and the like, but even if anyone could acquire the meaning of the word “knowledge” this way, we would only be replacing our existing word “knowledge” with a word synonymous with being in a certain brain state. But if the connection between a brain state and a mental state like knowledge is accidental, then we

will never be able to know that there are no circumstances where the connection fails, nor will we be able to know that we have canvassed all the possible brain states that might correspond to knowledge. More importantly for Wittgenstein, the fact that different criteria are used for mental phenomena and brain states shows that we are really dealing with two different kinds of thing.

The passages in *Zettel* reinforce this line of thinking. The sequence of passages begins with a warning: “One of the most dangerous ideas for a philosophy is, oddly enough, that we think with our heads or in our heads” (Z §605). The subsequent passages not only explain why this is mistaken, but also diagnose why we tend to think this way. The nature of the mistake is the same as it was when it was thought that we could find physical criteria for determining if two people were saying the same thing silently to themselves. We do not think with our heads in that our heads are not a necessary component for thought to exist. Our thoughts could also be the thoughts of an artificial intelligence. Hence Wittgenstein writes, “Is thinking a specific *organic* process of the mind, so to speak—as it were chewing and digesting in the mind? Can we replace it by an inorganic process that fulfills the same end, as it were use a prosthetic apparatus for thinking?” (ibid. §607). It may even be the case that now physical events are regularly associated with psychological events: “If I talk or write there is, I assume, a system of impulses going out from my brain and correlated with my spoken or written thoughts. But why should the *system* continue further in the direction of the centre? Why should this order not proceed, so to speak, out of chaos?” (ibid. §608).

The reason we are tempted to think that there is necessarily more going on here is that we are afraid that we will be unable to give psychological causes for physical events: “The prejudice in favor of psycho-physical parallelism is a fruit of primitive interpretations of our concepts. For if one allows a causality between psychological phenomena which is not mediated physiologically, one thinks one is professing belief in a gaseous mental entity” (ibid. §611). This is explicitly the line of thinking Jerry Fodor engages in all the time. His book *Psychosemantics* (1987) begins by asserting that the simplest explanation of the complex behavior of a cat involves entities like beliefs and desires. But if beliefs and desires are used to explain behavior, they must be related causally to that behavior, and if beliefs and desires have causal powers, they must be material, “because whatever has causal powers is ipso facto material” (ibid., x). This way of thinking is mistaken because the explanation that appeals to mental phenomena is of an entirely different kind than explanations that appeal to physical phenomena, with different criteria for correctness.

The upshot of this discussion, then, is that brain states can never be used as criteria for mental states. This is actually a specific instance of a more general problem, the impossibility of explaining normative phenomena empirically. The trait that unifies all the entities that we are saying cannot correspond to physical states is their *intentionality*. When one knows something, reads something, or says something to oneself quietly, one is engaged in an action directed toward something. The reason the physical states in question could not serve as criteria for these intentional acts is that they were unrelated to the intentional object. When one leaves out the intentional object of a mental state, one

also leaves out an important way in which mental states are normatively governed. Because the mental states in question are directed toward objects, they can also fail to describe those objects accurately. If I think or read something to myself, I can think or read something true or something false. Knowledge, of course, is assumed to be true, but this is still a normative constraint on it. The mental states in question are thus all normatively governed in a way that cannot be captured by the brain states that are supposed to correspond to them.

The hybrid nativist/apriorist program is compatible with Wittgenstein's ideas about mind and brain because it never conflates causal and intentional explanation. The model that the dispositionalist view is based on has three basic components. The first is Stanislas Dehaene's "triple-code" model of numerical processing. This is essentially a functional model of the mind, under which numerical information can be processed through either auditory, visual, or analog channels. There is also an attempt by Dehaene and his colleague Laurent Cohen to identify the parts of the brain that subserve these channels. Finally, there is a connectionist mock-up developed by Dehaene and Jean-Pierre Changeux. This mock-up consisted of a numerosity detector, which could identify the number of objects in a visual or auditory field, up to five objects, while discounting information about their size and location. This mechanism alone contains no ordinal information; however, Changeux and Dehaene add to it a complex that will learn to order numbers given basic input. The hybrid nativist/apriorist view accepts this model and adds the following claim: if the connectionist mock-up at all mimics the way humans actually

acquire mathematical knowledge, mathematical knowledge is dispositionally innate and negatively a priori.

None of these elements conflicts with Wittgenstein's view of mind and brain. When Wittgenstein warned us against providing physical explanations of mental phenomena, he reminded us of two ways of looking at a machine: one could think of its physical state, or one could think of its functioning. Problems only arose when the former was used as criteria for the latter. The elements of the hybrid position stick to the proper levels. The triple-code model is a functional analysis of the mind. It is no more problematic than a flow chart describing the functioning of a company (albeit a flow chart that was drawn by someone outside the company looking in). The triple-code model would only be problematic if it were taken to be more than this kind of flow chart. Identifying the parts of the brain subserved by the various aspects of the triple-code model would be tendentious if "subserved" were defined in problematic fashion. If, for instance, it was said that these regions of the brain were always necessary for an organism to engage in anything we would call mathematics, then a brain state would be invoked as a criterion for knowledge. But nothing of this sort is going on. Everyone these days is well aware of the multiple realizeability of mental functions. The connectionist mock-up is a machine that can be described either functionally or in terms of its physical states. Descriptions of connectionist machines often slide back and forth between the two, but this does not necessarily pose a problem.

Everything I have said so far is to be expected. Wittgenstein's views of the mind would be massively unreasonable if they challenged ordinary empirical research

programs like Dehaene's. The problems for such programs come up when they attempt to draw large philosophical conclusions about concepts and knowledge. (This is where the representationalist program ran aground.) Under the hybrid position, mathematical capacities are said to be the product of a disposition present at birth. Although empirical evidence is marshaled for this thesis, it is itself not a causal explanation because dispositions are not causal entities. Causation is a relationship between events or bodies, and dispositions are neither of these. At no point is a purely causal explanation being given for intentional properties. One might object that a causal explanation is given because we use empirical evidence to conclude the presence of a disposition. But the evidence used here is no different than the evidence used for a statement like "He was disposed to act kindly toward animals ever since he saw a sentimental documentary on animal feelings." One might make such a statement because one observes that the person in question engaged in more kind acts toward animals after seeing the documentary than before. The evidence for the disposition to mathematical knowledge is the same: it consists mostly of observations of behavior of people over the span of the human life.

So granted that the hybrid model is compatible with his views on the mind and the brain, what about the real thesis of this dissertation, the idea that mathematical statements are expressions of norms? The hybrid model complies with Wittgenstein's ideas about mind and brain by keeping causal and intentional explanation separate. So the only place where there might be a conflict with the thesis that mathematical statements are expressions of norms will therefore be the intentional parts of the model. But here I can see no difficulties. There is no problem with saying that mathematical knowledge is the

product of dispositions present at birth. These dispositions could easily be said to be dispositions to explain the rules one follows in dealing with objects a certain way. Nor is there a problem with saying that the mind processes mathematical information using three codes. These codes can carry information about norms as easily as they carry information about the empirical world. Finally, there is no trouble with connectionist models of how the brain actually carries out mathematics. Wittgenstein's thesis is a thesis about what the brain is doing, not how it does it.

We introduced the apriorist and nativist views because they could as easily explain the data presented in Chapter 2 as Wittgenstein could. We have now shown, however, that as substantial doctrines, nativism and apriorism fail. This makes Wittgenstein's normative hypothesis attractive again. Before anyone is likely to accept it, however, many things need to be clarified about it. In particular, we need a better understanding of how the objectivity of mathematics is possible. This is the task of the next chapter.

Numbers as Normative Facts

Introduction

The purpose of this chapter is to elaborate on exactly what it means to say that mathematical statements are expressions of norms. The focus will be on why mathematical statements are objective, contrary to common expectations regarding normative statements. When I presented some of the material from this dissertation in public for the first time, at a talk at Texas Tech University, one of the first questions I got was, “So does this mean that we should abandon our current training program for mathematicians and begin teaching them rhetoric?” The answer was “Of course not,” but exactly why I was entitled to give this answer turns out to be a fairly complicated matter. There are actually two worries that might be expressed here. The first is about objectivity of mathematics in the sense that the truth of mathematical statements is independent of our investigations. To dispel this worry, we will look at the social nature of norms, which will allow us to talk about the objectivity of expressions of norms. The second worry is about objectivity in the sense that is opposed to cultural relativism. Am I advocating a view that lets us dismiss mathematics as a European male construction? To allay this fear, I will argue that mathematical language is in fact compatible with all languages that talk of objects. My attempts to allay these two apprehensions may seem to be at cross purposes. The objectivity of norms in general depends on the possibility of rightful dissent from a universally held view, while the universality of mathematical language

seems to rule out such dissent. Complicating matters further is the fact that in Chapter 2 we appealed to the impossibility of universal error in mathematics as evidence for the thesis that mathematical statements are expressions of norms. These conflicts will be reconciled, however, when we understand what dissent means in different situations.

The investigations of this chapter will take us well beyond the bounds of what Wittgenstein considered legitimate philosophy. Wittgenstein thought that the philosophy of mathematics would end once we established that mathematical statements are expressions of norms. Any philosophical puzzlement one can have comes from misunderstanding the grammar of certain phrases. Once the grammar is clarified there is nothing more to say. The belief that mathematical statements were descriptions of the world was a grammatical confusion that led to a misguided debate over whether these descriptions were innate or learned, empirical or a priori. Once Wittgenstein had clarified the grammar of number, showing that mathematical statements are expressions of norms, this confusion ends. We are now free to go about more productive work, unburdened of the stupefaction philosophy brings.

Wittgenstein never justified the claim that all philosophical questions come from grammatical confusion. Nor could he, if he wanted to be true to his desire to present no general philosophical theses. Instead Wittgenstein worked by example, showing how individual philosophic problems disappear when language is understood properly. This chapter can be seen as a way of responding in kind, addressing a set of related problems, the problems of objectivity, that are not dispelled by clarifying the grammar of number. In fact they are raised by it. This chapter can also be seen as the third stage of the

dialectic begun by the earlier chapters. Having seen a thesis and an antithesis, we now get a synthesis—an explanation of what positive ideas come out of the battle that came before.

This chapter will proceed in three sections. The first will explain the social nature of norms, which will provide a foundation for the sections to come. The second section will address the objectivity of expressions of norms in general, reconciling the possibility of rightful dissent from any norm with the impossibility of universal error in mathematics described in Chapter 2. The third section will address specifically the objectivity of mathematical norms and the relationship between mathematical norms and object norms. Because this discussion goes beyond what Wittgenstein felt the role of philosophy should be, we will not be able to draw heavily on his work. In fact, the only claim about Wittgenstein I will make in this chapter is that he is *not* able to resolve the issue of the sociality of normativity. Instead my inspiration will be the work of Robert Brandom, who has done more than anyone else to present a positive account of language where the normativity of rules is irreducible.

Are Norms Inherently Social?

Throughout the *Philosophical Investigations* Wittgenstein hints at the idea that one cannot follow a rule by oneself. He seems clearly tempted by the idea that one can only make sense of a rule if one sees it against the background of a community that also follows the rule. Wittgenstein also just as clearly seems to deliberately back away from actually saying this. In a passage that has become the focus of controversy he writes,

Is what we call “obeying a rule” something that it would be possible for only *one* man to do, and to do only *once* in his life?—This is of course a note on the grammar of the expression “to obey a rule”

It is not possible that there should have been only one occasion on which someone obeyed a rule. It is not possible that there should have been only one occasion on which a report was made an order given or understood; and so on.—To obey a rule, to make a report, to give an order, to play a game of chess, are *customs* (uses, institutions). (*PI* §199)

He asks if a rule can only be obeyed at one time and by one person, but he only answers the first question. A rule must be obeyed on many occasions, but must it be obeyed by many people? The remarks about customs and institutions imply that the answer is yes, but why is he withdrawing here? In part he is probably motivated by his desire to leave as much work as possible to the reader: “What the reader can do, leave to the reader” (*CV*, p. 77). Unfortunately his reticence has led to a great deal of argument, both about what he was trying to say and about what he was actually entitled to infer from his arguments. Many argue that Wittgenstein definitely felt that rules were only intelligible in the background of a community, and moreover that the community is what resolves the two central Wittgensteinian problems we looked at in Chapter Two: the regress of interpretation and the gerrymandering argument. Others argue that the community view of rule following makes precisely the mistakes that the regress and gerrymandering arguments warned against.

I will argue for a position somewhere between the typical communitarian view of norms and the typical individualist view. Often the debate is framed as being between a communitarian view that says that all norms must be understood against a background of a community of actual humans and an individualist view that all norms can be understood in terms of a single individual’s relationship to autonomous grammar. As I see it,

Wittgenstein's arguments show that rule following must be understood against the background of a practice, and that this practice must comprise a multitude of judgements. However, the judges may be actual persons or aspects of the same person—time slices, impulses that may be in conflict, etc. The only necessary feature of a practice is its temporal extension, not its social extension. As we shall see, part of the reason this has been missed is that the significance of the temporal dimension has been misunderstood. Nevertheless, my stance is quite different than the stance taken by individualists like Gordon Baker and P. M. S. Hacker because I believe that *some* norms must be understood against the background of a social practice. Indeed, all of the important human norms require this background. For this reason, I will characterize my viewpoint as *weakly social* or *weakly communitarian*. In staking out this position, I will make no pretense about being true to Wittgenstein's views. I am instead stating the conclusion that I believe actually follows from the arguments he has given.

The secondary literature on the role of the social in Wittgenstein's thought is voluminous, largely autonomous from Wittgenstein's own work, and can easily be said to have spawned its own, tertiary literature. What follows should only be thought of as giving a rough outline of the debate. Those who advocate the social view of norms can basically be divided into two groups. The first group takes Wittgenstein's rule-following arguments to have demonstrated some kind of skeptical or antirealist conclusion about meaning. The community enters into these arguments as a kind of second-best source of meaning, something we resort to when we find that meaning does not have the kind of existence we wanted. The second group says that Wittgenstein has no problem with

saying that there are facts about meaning, and that social norms are real in the most robust sense possible. I will call these the skeptical and antiskeptical camps, respectively. Individualist interpreters, on the other hand, are all antiskeptical. In what follows I will present and critique the views of representative thinkers from these three camps. My own position will then come out in the course of this critique.

The locus classicus for the skeptical collectivist view is Saul Kripke's monograph *Wittgenstein on Rules and Private Language* (1982), based on lectures he had been giving since 1976. There are many similar interpretations of Wittgenstein, some antedating the publication of Kripke's work, including Fogelin (1976) and Wright (1980), but I will take Kripke as the exemplar of this genre. Lucid and compelling, Kripke's book helped shift the focus of Wittgenstein scholarship away from the private language argument to the remarks on rule following preceding it. These passages are now rightly regarded by most everyone as part of the core of the *Philosophical Investigations*. Kripke's book also, I think, attracted people who otherwise would not have been interested to Wittgenstein's work because it presented a linear argument with explicit roots in well-known arguments from David Hume. However, Kripke made no pretense of presenting Wittgenstein in Wittgenstein's own terms, or even presenting an argument that is fair to Wittgenstein in all of his moods. But he wasn't presenting his own views, either. He offered instead "Wittgenstein's argument as it struck Kripke" (1982, 5).

Wittgenstein struck Kripke as presenting one argument, encompassing both the regress and gerrymandering arguments, and placed in the mouth of a "bizarre skeptic" who seeks to convince us that "there can be no such thing as meaning anything by any

word” (ibid., 8, 55). Suppose that I had never added numbers larger than 57 before. The skeptic argues that if I use the symbol “+” in accordance with my previous intentions, I will say that $68 + 57 = 5$. This is because, in the past, I was not using “+” to mean “plus;” I was actually using it to mean “quus,” a function that always equals 5 when one of the operands is over 57. The skeptic considers a gamut of things about my history that might show that this was not my intention, and dismisses them all. But if there is nothing about me in the past that can fix my meaning, then nothing about me in the present can fix it either. It is then concluded that there is no meaning: “any present intention can be interpreted so as to accord with anything we choose to do” (ibid., 55).

In the face of this skeptical argument, Kripke’s Wittgenstein offers a “skeptical solution” similar to the skeptical solution Hume offered to the problem of induction. A skeptical solution is one that accepts the conclusion of a skeptical argument, but finds a way to live with it. Skeptical arguments are threatening, because they remove the justification for our ordinary beliefs and practices. We typically talk about people *meaning* the things that they say, and the skeptical argument Kripke sees in Wittgenstein removes our justification for doing this. The skeptical solution saves us by showing that our ordinary practices “did not require the justification that the skeptic has shown to be untenable” (ibid., 66). In this case the task is to show that we can still talk about people meaning what they say, even if any action can be made to accord with that meaning.

The key to the skeptical solution, according to Kripke, is the shift from explaining language in terms of the truth conditions of assertions to explaining language in terms of two questions that can be asked of any linguistic act: “When is it appropriate to make this

linguistic act” (the assertability conditions) and “What is the purpose of making this linguistic act.” The skeptical solution is found when we apply this new approach to language back to statements such as “*x* means *y* by *z*.” The skeptic has shown that we can not find truth conditions for such statements. However, we can still easily talk about the assertability conditions for “*x* means *y* by *z*.” I say such things of myself when I feel confident that I am applying a rule in accordance with my past intentions. I am confident that I meant plus, not quus, by “+,” and am entitled to assert as much. (This is Kripke’s gloss on Wittgenstein’s talk of acting without justification, but not without right.) We also look to our own inclinations in judging the meaning of others: “*Smith* will judge Jones to mean addition by ‘plus’ only if he judges that Jones’s answers to particular additions problems agree with those *he* is inclined to give, or if they occasionally disagree, he can interpret Jones as at least following the proper procedure” (ibid., 91). This part of individual practice is what allows us to say that someone thinks they are following a rule, when they actually are not. The social enters Wittgenstein’s thought, according to Kripke, when it comes time to assess the usefulness of asserting statements about meaning. “If we were reduced to a babble of disagreement, with Smith and Jones asserting of each other that they are following the rule wrongly, while others disagreed with both and with each other, there would be little point to the practice just described” (ibid.). What saves us from anarchy is society’s practice of granting individuals status as autonomous rule followers if their judgements match the judgments of the community most of the time, or can at least be understood as following the same procedures as the community. On this view, the correctness or incorrectness of an action is indirectly the

consequence of the community sanctioning that action. Because the community only recognizes an individual as having mastered a rule if the individual agrees with the community most of the time, or can at least be understood as following the same procedures as the community, an action will be correct if and only if it earns the consensus of the community.

Many objections have been advanced from many quarters against the skeptical, communitarian view of rule following. One of the most often heard, but actually least interesting, is the protest that Wittgenstein would have nothing to do with skepticism. It is, of course, certainly true that Wittgenstein believed for his whole life that skeptical problems are pseudoproblems. As Gordon Baker and P. M. S. Hacker (1984, 5) point out, remarks to this effect can be found from Wittgenstein's earliest published writings (*NB*, p. 44) to his last (*OC* §495). But this point packs less punch than one would think, if only because it is something that Kripke acknowledged (1982, 63). More importantly, Kripke was right to go ahead with his reconstruction of Wittgenstein's argument in the *Philosophical Investigations*, even though it contradicted one of Wittgenstein's core beliefs. The fact is, Wittgenstein was a man of deep contradictions. Any attempt to extract an interesting argument from his work is going to do it injustice somewhere. Very likely, the injustice will be done to his metaphilosophical views, since some of the most profound contradictions in his life came out of his attempts to resist philosophy.

A far more important objection that comes up often focuses on whether the community view of rule following can solve the skeptical problems once they have been raised. A common version of this argument comes from Simon Blackburn (1984), who

directs it against Kripke. Blackburn basically puts a dilemma to Kripke: either being a community-approved rule follower is a part of the truth conditions for the predicate “___ follows a rule correctly” or it is merely a part of its assertability conditions. If it is a part of the assertability conditions, then any solution community approval offers for the skeptical paradox can also be offered by the individual. The individual, viewed as a collection of time slices, can as easily grant herself status as a bona fide rule follower by looking at her past selves and seeing if their actions agree with each other. She may even deny herself that status if she finds that her actions do not coalesce into a coherent practice. If, on the other hand, being a community-approved rule follower is a part of the truth conditions for “___ follows a rule correctly,” then rule-following considerations show that the individual cannot grant herself status as a rule follower. But then again, neither can the community. The community cannot look at their previous actions and decide whether they were following a normal rule or a bent rule. They could be a “thoroughly Goodmanned community in which people take explanations and exposure to small samples—yesterday’s applications—in different ways” (ibid., 294). Analogues to Blackburn’s argument, or elements of it, crop up in the arguments of others. One half of Blackburn’s dilemma is mentioned by Baker and Hacker in the course of their vituperative attack on Kripke, *Skepticism Rules and Language*, “Given that no one previously ever added 57 and 68, how do we know that our present community-wide inclination to answer ‘125’ accords with what we previously meant by ‘plus’, i.e., with what we would have been inclined to say, had we previously been asked what $57 + 68$ is?” (1984, 37). A similar point is made by Colin McGinn: “It seems to me that

essentially the same problem of transtemporal normativeness comes up with respect to the whole community: for can we not ask what justifies our assumption, as a community, that yesterday *we* meant the same by our words as we do today?" (1984, 188).

Blackburn's argument shows that the community has no more resources to solve the skeptical paradox that has been offered than the individual. This, however, only challenges the skeptical solution, it has not rid us of the skeptical paradox. Following McGinn, I would like to suggest a "straight" (nonskeptical) solution to the skeptical paradox. In his book *Wittgenstein on Meaning* (1984) McGinn suggests that one can adequately reply directly to the skeptic if one believes that meaning is irreducible. Since the irreducibility of meaning is a consequence of Wittgenstein's rule-following arguments as I outlined them in Chapter 2, this is a very natural line of argument for me to endorse. Let's look at McGinn's argument.

When Kripke develops his skeptical paradox, he considers many ways one might justify the claim that one has been using "+" to denote "plus" rather than "quus." All the candidates he considers seriously, however, reduce meaning to something else—mental images, an internalized set of directions, dispositions, qualia, etc. But if we say that meaning something is an irreducible fact about an individual, we have a natural reply to Kripke. Why can't we give the truth conditions for a sentence about meaning "by simply *re-using* that sentence, frankly admitting that no other specification of truth conditions is available—precisely because semantic statements cannot be *reduced* to non-semantic ones" (McGinn 1984, 151). Kripke does consider the possibility that someone might reply to the skeptic in this manner, but he does so only briefly, and his remarks are not

compelling. “Such a move may in a sense be irrefutable, and if it is taken in an appropriate way Wittgenstein may even accept it. But it seems desperate: it leaves the nature of this postulated primitive state—the primitive state of ‘meaning addition by “plus”’—completely mysterious” (Kripke 1982, 51). It is mysterious for two reasons. First, it cannot be a kind of irreducible qualia, because he has already argued that qualia cannot fix meaning. But if this primitive state is not a qualia, how are we supposed to be aware of it? The second source of mystery is the fact that meaning is supposed to be infinite, but our minds are finite. McGinn has replies to both of these points. In reply to the first claim, he points out that our first-person knowledge of many psychological states, such as our knowledge of our beliefs, thoughts, intentions, and hopes, does not come with an associated qualia. As for the second argument, McGinn maintains that this is a problem about infinity, not a problem about meaning as such, moreover, it is not a problem that applies only to meaning. “Consider, by way of analogy, our beliefs and desires: do they not have ‘indefinitely many’ (normative and casual) consequences?” (McGinn 1984, 163). McGinn also has an openly ad hominem argument on this issue: Kripke himself (1972), when not acting as a filter for Wittgenstein, has endorsed a nonreductive account of reference, which would surely be mysterious if the nonreductive account of meaning is mysterious.

McGinn’s argument is a natural one for me to endorse, because the irreducibility of meaning follows naturally from the claim that normativity is irreducible, which I attributed to Wittgenstein in Chapter 2. There I claimed that the rule-following arguments were designed to show that normativity is irreducible. Following Brandom (1994) and

Meredith Williams (1991/1999), I divided Wittgenstein's ideas on rule following into two arguments, the regress and the gerrymandering arguments. Each was designed to show that nothing can stand apart from the applications of a rule, determining which are correct and which are incorrect. The regress argument emphasized the multiplicity of guides that might be invoked to justify an action, whereas the gerrymandering argument emphasized the multiplicity of ways a single guide could be interpreted, but in each case the outcome was that the correctness of an action could not be separated from the action itself. By leaving normativity intact as a real feature of the world, this argument contrasts sharply with Kripke's argument. More importantly, it opens up the possibility of using McGinn's arguments to reply to Kripke's skeptic. If normativity is irreducible and meaning has a normative component, then meaning something is an irreducible property of the individual. If meaning is irreducible, then McGinn's argument follows.

So we have ruled out one route to a social conception of normativity. We do not need the community view to save us from a kind of skepticism. First, the community view cannot save us from such a skepticism. Second, the skepticism itself is unwarranted because it ignores the possibility that meaning is irreducible. Other interpreters have argued, however, that Wittgenstein endorsed a community view of normativity, without ever proposing a skeptical problem. Williams has propounded such an interpretation in a series of articles (1983/1999, 1990/1999, 1991/1999, 1994b/1999). For Williams, many aspects of Wittgenstein's thought lead us to the conclusion that norms must be social, not just his rule-following arguments. The considerations which, on Williams' view, lead to a social conception of norms tend to blend and overlap—perhaps an inevitable

consequence of being true to Wittgenstein's actual thought. For expository purposes, I will divide them into three categories. First, there is Wittgenstein's critique of ostensive definition, which for Williams gives the private language argument its force. Second, there are the kind of rule-following considerations Kripke focuses on. Finally, there is the nature of language learning.

Kripke took Wittgenstein's rule-following arguments to be the heart of the *Philosophical Investigations*. But the *Investigations* actually opens with Wittgenstein's famous discussion of ostensive definition (*PI* §§1–38, esp. §§28–38). For Williams, this discussion plays a key role in Wittgenstein's arguments on the sociality of norms, including a much neglected role in the private language argument (1983/1999; see also 1990/1999, 147–53 and 1994b/1999, 190–96). The point of the discussion of ostensive definition in the beginning of the *Philosophical Investigations* is that we cannot teach a first language by acts of ostension alone. Ostension is actually a very sophisticated act; in order to know the use of the word uttered while pointing, and identify the exact object being pointed at, one must already be a competent language user. If one is not already a competent language user, one will not be able to distinguish the name of an object from a command to fetch the object. Nor will one be able to tell whether it is the object that is being named, or its color, or the number of objects in its vicinity. In truth, language acquisition requires two components. The first is a process that Wittgenstein calls "ostensive teaching" (*PI* §6). As Williams sees it, this kind of ostension is prelinguistic and within the capacity of animals. All it does is set up a bare association between signs

and objects. Full language ability only comes when the language learner understands the use of these signs. Passing on use is thus the second essential component of learning.

According to Williams, the private language argument is simply an application of this general critique of ostension to the idea of private ostension. The private language argument asks us to imagine a person who tries to create a language only she can understand, a language whose words refer to the sensations that only she has. For the private linguist to succeed, she must be able to do two things. First, she must be able to point mentally to one of her sensations to designate it as the referent of a word in her language. Second, she must be able to go on to use that word consistently. The early commentators on the private language argument have assumed that Wittgenstein's critique is directed at the second requirement of a private language. (The assumption is natural enough, given that Wittgenstein does talk about the need for a "criterion of correctness" [*PI* §258].) However, Williams believes that Wittgenstein's real target was the first requirement of a private language, the ability to make private, ostensive definitions. An isolated individual simply does not have the resources to use ostension. Ordinary ostensive definition requires a public language in order to know the use of the word uttered in conjunction with the object and the exact identity of the object pointed at. Ostensive teaching, on the other hand, also cannot function in private—one cannot be an ostensive autodidact. Ostensive teaching does not work unless use is also imparted, but use implies a public sphere. Therefore ordinary ostensive teaching "also presupposes a public language, although the child does not know it" (Williams 1983/1999, 21). But if the private linguist cannot establish definitions for her words, then any linguistic rules she

develops will be contentless. This is where Wittgenstein's talk of criteria of correctness comes in. Typically, it is thought that there can be no standard of correctness for the private linguist because of the lack of a public check on her applications of rules. However, Williams argues that the real reason the private linguist lacks a real standard of correctness is that any standard she uses will be contentless because of her inability to use ostensive definitions. The private language argument and the critique of discussion of ostensive definition thus give us our first point of contact between normativity and community. In order to know how to use words correctly, I must be a part of a community of language users. More broadly, I cannot give myself a rule with any content without the background of a linguistic community.

The social nature of norms is also manifested in the sorts of rule-following considerations Kripke focused on, and which we discussed in Chapter 2. According to Williams, a rule cannot be followed by only one person because one cannot make sense of right and wrong action unless there is a background of community agreement. (1991/1999; see also 1994b/1999, 197–206). However, community agreement is not constitutive of the correctness or incorrectness of an action, as it winds up being for Kripke. Nor is the role of the community to provide a check on individual action, as is often thought by interpreters of the private language argument. The easiest way to see why Williams believes this is to look at her treatment of a standard example in this debate, the case of Robinson Crusoe decorating his house. Suppose Crusoe paints a design on the walls of his house consisting of many repetitions of a master pattern. Why can't we say that Crusoe is following a rule? Well, to be sure he is intentionally following

a rule, we need to know that he is correcting himself if he deviates from the master pattern. So let's suppose further that he checks the repetitions against the master pattern and when the two fail to match, he makes an annoyed gesture and changes the repetition to match the pattern. Now can we say that Crusoe is following a rule? According to Williams, we can only do so by projecting our own community into the example: "Insofar as our isolated individual's behavior counts as being corrective, it is only in virtue of his behavior being like our own. The only standard available for what counts as corrective behavior, and so *is* corrective behavior, are the paradigms of correcting that inform our practices" (Williams 1991/1999, 173). We think Crusoe is correcting himself only because what he does resembles what our community does when it corrects himself. Without this resource, we have no way of making sense of what Crusoe is doing. Indeed, all sorts of bent-rule hypotheses suggest themselves. Perhaps the pattern Crusoe was following called for systematic changes in the repetitions, like the strange rule for adding two that leads one to add four after one reaches 1000. Perhaps the repetition Crusoe originally drew was the correct one, and "Crusoe's gesture that we took as a sign of annoyance was rather a sign of rebelling, his examination of the master pattern and his subsequent behavior a rejection of that master pattern for another" (ibid., 174). Therefore to see Crusoe as following a rule, we must place him in a social context. This, it should be emphasized, is a much stronger result than the conclusion of the private language argument. The private language argument eliminated the possibility of *necessarily private* rules. If one person invents a language, it must be at least possible for someone

else to understand it. Here Williams is saying that all languages are *necessarily public*. We cannot understand a language until it is placed in a public context.

But what, exactly, does this public context contribute? How does it let us see what rule Crusoe is following? Williams emphasizes that the community does not play the role of a judge, as it did for Kripke: “The very emphasis that commentators have placed on *corrective* behavior,” she writes, “is out of place. Error and correction are very much the exception. Checking, whether by others or oneself, simply doesn’t feature prominently in the exercise of a practice, except in the case of the learner.” Instead it is the *agreement* of our practices that lets us talk of right and wrong. Because the members of a community naturally react the same way to novel situations, they provide a background against which we can see an action as right or wrong. But this does not mean that community agreement constitutes right and wrong. The correct way of continuing a series is not *defined* as the way the community would continue it: “I certainly don’t mean, when I assert that $2 + 2 = 4$, that the majority answers “4” when asked, “what is $2 + 2 = 4$?” Nonetheless, without this agreement $2 + 2 = 4$ would cease to be meaningful” (1994b/1999, 202). Nor could I appeal to community consensus to *justify* the way to continue a series: “What is indispensable for correct, or appropriate, judgement and action is that there *is* concord, *not* that each individual justifies his (or anyone else’s) judgement and action by appeal to its harmony with the judgement of others” (1991/1999, 176). Learning is again essential here. Because normativity requires a background of social practice, there is no way to acquire the ability to follow a rule without being initiated into a community of rule followers. “Given the social foundation of norms, acquiring an understanding of rules and

other norms requires assimilation into community practices. The point is not just that there is a de facto causal link between the training of a novice and assimilation into a practice, though this is obviously true. The point is rather that there is no other way to explain the acquisition of *normative* competencies” (1994b/1999, 203).

The discussions of rule following and of ostension and the private language argument both talked about learning to follow rules. Learning can actually be thought of as a separate point of contact between normativity and the social for Williams. In her view, the process of learning is actually constitutive of what is learned. Williams argues this in her essay “The Philosophical Significance of Learning in the Later Wittgenstein” (1994b/1999), by a close analysis of Book VI of the *Remarks on the Foundations of Mathematics* and Wittgenstein’s metaphorical description of learning a rule as traveling in a circle. The upshot of the preceding arguments is that all rules depend on a background of practice. The rule would not be what it is without that background of practice. A practice cannot be reduced to a set of rules because it is what makes rules possible. Attempting to reduce a practice to a set of rules leads to rule-following paradoxes. The only way one can understand a practice is to be trained into it. Being trained in a practice means acquiring a sense of the obvious. The members agree unthinkingly in their actions because they intuitively feel the same sorts of things to be obvious. Wittgenstein’s metaphor of traveling in a circle is meant to capture how one learns what is obvious. The student first performs an action without understanding how it should come out. At first, this is done as a test, to see if the pupil can get the answer right. During the process of learning, the teacher constrains the possible actions of the pupil,

until it seems obvious to the pupil that only one outcome is possible. “The activity of testing itself is transformed into one in which getting a particular outcome is essential to the process” (ibid., 210). Once this result dependence, as I have been calling it, is impressed on the pupil, the same action takes on a different meaning. It is no longer a test, but an exemplar of how a process should work. But if a rule would not be what it is without the background of practice, and the only way to understand a practice is to enter into it through this circular process, then the process of learning is constitutive of the content of a rule.

My disagreement with Williams’ view is much less severe than with Kripke. Nevertheless, it is important. I agree that part of the role of a practice is to provide a background of agreement that allows us to talk about correct and incorrect. Williams is also certainly correct to point out that we cannot appeal to this agreement to justify an action, nor is this agreement part of the definition of right action. This practice, however, can exist entirely within an individual. Two remarks from Blackburn (1984) can help us see why this is plausible. The first, already noted, is that an individual can easily be seen as a community of time slices. The second is a part of Blackburn’s critique of the private language argument. Blackburn claims that Wittgenstein has not given the private linguist access to all the internal resources he would actually have: “In the usual scenario, the correctness or incorrectness of the private linguist’s classification is given no consequence at all. It has no use. He writes in his diary, and so far as we are told, forgets it. So when LW imagines a use made of the report (e.g., to indicate the rise of the manometer) he immediately hypothesizes a public use. He thereby skips the intermediate

case where the classification is given a putative private use. It fits into a project—a practice or technique—of ordering the expectation of recurrence of sensation, with an aim at prediction, explanation, systematization, or simple maximizing of desirable sensation” (ibid., 299–300). Within an individual person, we can identify other individuals—time slices of a person, conflicting impulses within a person, etc. These virtual individuals can agree or disagree in their judgements, establish goals for themselves, and do well or badly at achieving those goals. It makes sense to talk about *use* in this internal economy. This kind of internal practice will not be *necessarily* private; others can come to understand it. The existence of an internal practice, therefore, does not conflict with the conclusion of the private language argument as it is commonly understood.⁴⁶ However, it can also be understood in its own terms, without reference to an interpersonal practice, and therefore does contradict the stronger results of Williams’ thoughts on rule following. It will no doubt be objected that this internal, faux community can only exist as a model of the external, real community. But before I turn to this complaint, I want to look at exactly how an internal practice can allow for all of the phenomena Williams discusses: ostensive definition, rule following, and learning.

I want to start with Williams’ discussion of rule following because it was there that we first saw the claim that not only were there no necessarily private rules, all rules were necessarily public. They required the background of a public practice to be understood. I presented Williams as offering a dilemma in the Robinson Crusoe example. On the one

⁴⁶ I will use the phrase *internal practice* to refer to the kind of entity I think is possible in order to distinguish it from a *private language*, which Wittgenstein shows is impossible.

hand, we can project our community onto Crusoe, in which case he clearly follows a rule, but has a social context. On the other hand, we can consider him in isolation, in which case he could be doing anything. But what of Crusoe's own perspective on what he is doing? Isn't there a third option: using Crusoe's own practice to explain which of the many possible rules he is following? Williams considers this possibility: "What about Robinson Crusoe himself? ... From his own point of view he engages in a repetitive behavior derived from a series of dots and dashes initially drawn on the wall, 'derived' in the sense that he reproduces the same series of dots and dashes as produced in the original sequence" (1991/1999, 174). But this is hardly how Crusoe would describe his own practice. He would say something like, "Well, I've got a lot of time on my hands, and this pattern means something to me, because I like its symmetry and I find the uniformity soothing," etc. This description indicates that there is an internal practice behind Crusoe's behavior. In many moods, at many times, he sees the same things as relevant about the pattern. He shares a sense of the obvious with himself over time. Moreover this agreement in unthinking judgment allows him to give a sense of purpose to what he does. It soothes his mind and fills his day.

It might be objected at this point that I am falling back on projecting my own community standards on Crusoe. After all, I am letting him speak in English, and express judgments in terms we all share, like "symmetry" and "soothing." But really I am not doing anything with Crusoe's internal practice that we would not have to do to understand *any* foreign practice, individual or communal. In order to understand the practices of an alien culture, I also need to assimilate their culture to my own. But this

does not lead me to say that all rule following requires a background of *my* culture's practices. Understanding rule following requires the background of some practice, and we always start with the practice we are a part of, but this does not mean that any particular kind of practice is privileged. To make this clear, I want to change the thought experiment so that we do not have to think about humans and their social nature. Imagine a species of space aliens that are as solitary as turtles. They meet only to mate and lay eggs in the sand, which they promptly abandon. If the environment is tricky enough, it is at least logically possible that these creatures could evolve into users of a contingently private language. In order to keep track of the population of their prey, for instance, they each establish a personal record keeping system. Now it would be perfectly possible for humans traveling to this planet to make sense of the idiolect of one of these creatures. To do so, we would begin by assimilating its practice to ours. This is simply good hermeneutic practice. The process would probably end with the human initiated into the alien practice, understanding it in its own terms. But the important thing to note is that the situation is perfectly symmetrical. To understand us, our solitary alien would begin by assimilating our social practice to its internal one. Perhaps it would say that all rules must be understood against the background of an internal practice!

Once we see how an internal practice can allow for rule following, it is easy to see how it can allow for internal ostensive definition and self-teaching. The thrust of Wittgenstein's remarks on ostensive definition was that it was not enough to set up a bare association between an object and a sign. The sign needs to have some *use*. The object, as well, needs to be a part of a practice, if we are to know exactly what is being ostended.

But if we allow for some complexity in the internal workings of an organism, both of these things are possible. Imagine our solitary alien life form learning to identify hunger in itself. On having the sensation of hunger, it makes a note to itself to remember the sensation, determine what causes it, and avoid that cause. As an isolated moment, this ceremony is entirely empty. There is no way to know what aspect of its sensory field the alien is pointing to. A social practice would allow the target to be specified, because all the members of the practice would share a sense of the obvious, and automatically pick up on the object of ostension in the same way. An internal practice can do the same thing. The alien, in all of its moods and moments, will find the same things significant about its sensation. This homogeneity will also let us talk about the alien giving a purpose to its ostension.

From here, it is not hard to explain how the alien could become an ostensive autodidact. Suppose the alien wants to teach itself to flee automatically at the sight of human colonists. Currently, when it sees these strange creatures, it does not know how to react. Sometimes it runs; sometimes it charges; sometimes it freezes. Divide the alien into two virtual persons: alien-in-repose who has decided to avoid the colonists, and alien-in-the-face-of-earthlings, who is panicked and reacts unpredictably. The more rational side of the alien can guide the actions of the panicky side, punishing itself for poor reactions, rewarding the correct reactions, until it reacts automatically to humans by fleeing. This reaction is more than a bare association, because it has a role in the internal economy of the alien. The alien-in-the-face-of-earthlings now shares a sense of the obvious with the alien-in-repose. It is now a part of the practice of avoiding humans.

The objection I alluded to earlier, that this kind of internal, faux community can only exist as a model of the real community will no doubt surface again here. Yes, it will be said, you can metaphorically divide a person up into subpersons, but this metaphoric use of the concept of community will always be parasitic on the existence of the real community. It is the real community that we first encounter—first we learn to talk, then we learn to talk to ourselves. More importantly, it is the real community that is required as a background for understanding norms.

There are actually two parts to this objection. The first is a point about the genesis of norms and practices. Humans first are initiated into a shared practice, and then learn to set up things that we can metaphorically call “internal practices.” This is basically an empirical point, and I will not dispute it. It may or may not be true that we model the internal community on the external community. Many in cognitive science, like Marvin Minsky, like to talk about the society of mind. It is entirely possible that we are innately disposed to develop into such a community. More importantly, even if it is true that we model an internal community on the external community, this is only a matter of the actual, contingent genesis of norms. Humans learn to follow norms by being initiated into a community practice, and later learn to set up such practices internally. The second part of the objection was the conceptual claim that all norms must be understood against the background of a *real* practice, by a *real* community. I agree that this happens to be the case for all the important human norms, but this is only a contingent feature of norms. A creature that derives its norms solely through an internal practice is a logical possibility. Such a creature would have to understand our norms against the background of its

internal practices, just as we would understand its norms against the background of our social ones.

The only essential feature of a practice is its temporal extension. This point has been missed, I think, because people have assimilated the temporal extension of a practice to the multiple applications of a rule. The *multiple applications thesis*, as McGinn (1984, 37) has dubbed it, says that to make sense of rules at all, at least some rules must be applied multiple times. But there is more to the temporal extension of practice than this. A practice requires a concurrence of automatic, unthinking, judgments. A fortiori, it requires multiple acts of judgement. Temporal extension is a necessary condition for this. An individual clearly must be extended over time to engage in multiple evaluations. A group must also be extended in time, because it takes time for one individual to react to another. Temporal extension is also necessary to talk about goals and purposes, which are needed if we are going to talk about use. On the other hand, I have argued that social extension is necessary for neither of these things. The crucial element in practice, and hence in norms and meaning, is time.

In case the account I am presenting comes across as too strongly individualist, I would like to take some time to contrast it with the hardcore individualism of Baker and Hacker. According to Baker and Hacker, the point of Wittgenstein's rule-following arguments was to show that there is an "internal relation" between a rule and its applications. Baker and Hacker are not always clear about what they mean by "internal relation"—they say so many things about the concept it is sometimes hard to identify its basic meaning. However, at bottom they are using the phrase in the standard

philosophical sense. Typically a relationship is called “internal” if being in the relationship is an essential property of one or both relata, and this is exactly what they say about internal relations: “A relation between two entities is internal only if it is inconceivable that those two entities should not stand in this relation” (1984, 107). At times, Baker and Hacker’s use of “internal relation” seems to conflict with this traditional meaning, but the appearance is illusory. For instance, they claim that “Wittgenstein repudiates the implication that any expression of an internal relation must be a necessary truth or a tautology” (ibid., 109). This might be taken to mean that internal relations do not involve the essence of the relata. However, the examples Baker and Hacker use to illustrate this idea show that all they are asserting is that the relata are not identical. So, for instance, they say that even though there is an internal relation between pain and pain behavior, either can exist without the other (ibid., 110).

The key feature of an internal relation for Baker and Hacker is that an internal relation cannot be explained in terms of a third thing: “an internal relation between two entities cannot be decomposed or analyzed into a pair of relations with some independent third entity” (ibid., 107). This is, in essence, what rules out social conceptions of rule following for Baker and Hacker. By claiming that one must understand norms against a background of a social practice, one is interposing a third entity between a rule and its application. Instead, rules and their applications meet in a “grammar” that is independent of all social considerations. Baker and Hacker describe grammar as “autonomous,” meaning that no fact about the world can be used to dispute or support a statement of grammar. Grammatical statements are “*ex officio* immune to impeachment. Nothing

counts as a legitimate challenge to a rule of grammar, and hence nothing qualifies as a justification of it” (ibid., 111). Grammar is autonomous because it is “antecedent to truth” (ibid., 99). Grammar “delimits the bounds of sense; hence any description of reality put forward to justify grammar presupposes the grammatical rules” (ibid.).

Baker and Hacker’s claim that nothing can mediate the relationship between a rule and its application contradicts my claim that a rule must be understood against the background of a practice that comprises many acts of judgement. The fact that I countenance virtual persons within an individual does not make me less of a communitarian from their perspective. I am still trying to split up an internal relation. This would be bad news for me, if Baker and Hacker had an effective critique of all those who attempt to split internal relations. Fortunately for me they do not. By rendering grammar completely autonomous from the empirical world, they also render it quite mysterious. As Williams has pointed out, Baker and Hacker run the risk of making grammar the very kind of “philosophical superlative” Wittgenstein warned against (1991/1999, 166). Like the an interpretation or the machine-as-symbol, autonomous grammar seems to encompass all the applications of a rule “in a flash.” It is an atemporal attempt to capture an essentially temporal phenomenon, normativity.

The Objectivity of the Expression of Norms

Although the practices that underlie norms are not necessarily social, as a matter of fact, all human practices are social practices. The goal of this section is to show how the practices that give rise to norms also allow them to be objective. To do this, I will adopt

Brandom's deontic scorekeeping model of practices and the norms they yield. Brandom's account is social, in that it requires multiple scorekeepers. However, these scorekeepers can be actual separate people, or time slices of an individual. It is therefore compatible with the weak social view I adopt. Brandom's account has two premises. First, rule following necessarily involves evaluation. One must not only act, but evaluate that act. (Brandom claims this is an adaptation of Kant's view that when rational beings follow a rule they are acting on their concept or representation of a rule.) Second, any account of rule following must allow for the possibility of following a rule incorrectly. These premises imply that the evaluation of an action with regard to a rule cannot be identical to following the rule. Evaluation cannot simply be a regular differential response to someone acting on a rule. Iron offers a regular differential response to humidity: it rusts. But we do not want to say that it is evaluating the humidity of its environment in the relevant sense. Evaluation requires something more. Brandom suggests that the additional criterion is the presence of sanctions. One would then count as evaluating actions if negative evaluations were accompanied by some sort of punishment and positive evaluation some sort of reward. Brandom emphasizes that sanctions themselves might be defined purely normatively. A community might punish an offender by beating her with sticks, or they might simply say that the offender does not have the right to attend a sacred feast.

The social enters Brandom's theory when it comes time to introduce objectivity into the evaluations surrounding language use. Two different kinds of objectivity are important here. First, our speech acts and our evaluations of our speech acts must be right

or wrong apart from anyone's saying they are right or wrong. Second, our utterances must be about *objects* that exist apart from our descriptions of them. Both of these sorts of objectivity are introduced through a model of the normative aspects of linguistic practice Brandom calls *deontic scorekeeping*. Under this model, every member of a linguistic community keeps score on themselves and every other member of the community, tracking what commitments they have made (for instance, what statements they are committed to the truth of based on the assertions they have made) and what further actions and utterances they are entitled to make (for instance, assertions they are authorized to make based on inferences from claims they are committed to). It is a structural feature of this system that any person can be wrong about what they are committed and entitled to. No matter how one person or group of people have evaluated an assertion, other evaluations are possible. This allows for objectivity in our first sense. The second sense of objectivity dealt with objects themselves. This sort of objectivity is allowed for through the use of *de re* ascriptions. Suppose someone, call her Sally, makes the assertion "Ben Franklin was a printer." Because Sally is committed to the proposition that Ben Franklin was a printer, she is also committed to the proposition that the inventor of bifocals was a printer, even if she does not know that Ben Franklin was the inventor of bifocals. When another scorekeeper attributes this commitment to Sally, he will be making a *de re* ascription, as opposed to a *de dicto* ascription. He is taking her assertion that "Ben Franklin was a printer" to be about the thing (*re*) named in the sentence as it exists, not as its is named (*dicto*). *De re* ascriptions give life to objects beyond any one person's beliefs about them because they allow for an object to enter one's set of

commitments under a description that one is not aware is true. De re ascriptions also require multiple scorekeepers. Someone else must exist to ascribe to Sally a belief about Ben Franklin under the description “the inventor of bifocals,” because she is not aware of this description. Note that here again the argument takes the form of an inference to the best explanation.

Brandom’s account links normativity and the social, but it does not reduce normativity to community agreement, as Kripke did. For Brandom the rule always exists apart from anyone’s evaluation of it. One might claim that Brandom’s account is only nonreductive in the way that Kripke’s account was nonreductive. Kripke said that he was not offering a community view of the truth of statements. Thus he was not saying that it is true that we follow a rule correctly when we agree with our fellows. He was merely offering an account of the conditions under which we will *say* that a person is following a rule correctly. Brandom’s deontic scorekeeping seems to work a similar way. It does not provide an account of when one is actually following a rule, it merely describes the way we attribute rule-following behavior to each other. There is a crucial difference, however, in the way Brandom and Kripke’s accounts are presented. Kripke presented his social view as the only solution to his skeptical paradox, allowing Blackburn to point out that there was a parallel individual account. Brandom does not present the social view of rule following as a solution to any paradoxes and is free to say that there are in fact truth conditions on following a rule, as long as the nature of truth conditions are properly

understood. Brandom, like Wittgenstein, adopts a roughly deflationary account of truth.

So, in essence, “ x means y by z ” is true when x means y by z .⁴⁷

The social entered Brandom’s account when it came time to explain how language can be used objectively—how it can be right or wrong and how it can refer to objects in the world. The concern for this section is how a specific kind of linguistic act, the expression of a norm, can be objective. Basically, this happens for expressions of norms the way it happens for any kind of linguistic act. Expressions of norms can take two forms. One can either say that a certain act is in keeping with an already accepted rule, or one can advocate the acceptance of a new rule. In terms of deontic scorekeeping, this means that one can either assert that commitment to a certain rule entails commitment to a certain act or judgement, or one can say the commitment to a certain rule or set of rules entails commitment to another rule or set of rules. Mathematical statements can take either form. The former sort of statement would be something like “If you believe that $2 + 2 = 4$, then you must believe that that there are four objects here.” Statements like this are a hybrid of pure mathematics and applied mathematics. The consequent of the entailment is an applied mathematical statement, in my terms, and as a result is descriptive. The antecedent of the entailment is a pure mathematical statement, and hence normative, and the claim that these propositions are related by entailment is also normative. The other sort of expression of a norm in mathematics would be something

⁴⁷ Actually Brandom endorses an anaphoric account of truth, under which “__ is true” is a prosentence forming operator. However “ x means y by z is true” is only a case of what is called “lazy anaphora,” so in this case there is no difference between Brandom’s account and the traditional disquotational account. (See Brandom, 1994, 285–305.)

like “If you believe these axioms you must believe these theorems.” This claim is all on the level of pure mathematics. The statement “ $2 + 2 = 4$ ” itself is also on this level. It says that if one is committed to a statement involving the term “ $2 + 2$ ” one is also committed to the same statement with “4” substituted for “ $2 + 2$ ” (and vice versa).

Both of these kinds of expressions of norms have both of the kinds of objectivity we outlined earlier. The norms governing language use were objective because it was possible for a statement to be right or wrong apart from anyone’s saying it was right and wrong, and because statements could be about an object that exists apart from anyone’s description of it. Both of these properties apply equally well to statements of norms of all types. Not only are statements about the empirical world true or false apart from anyone’s say so, statements about whether actions and judgements are entailed by norms, as well as whether norms entail other norms, are also true or false apart from anyone’s say so. These kinds of statements are going to be cashed out in terms of deontic scorekeeping. Under this system, it is perfectly possible for different people to attribute different sets of commitments to the same person, and hence to disagree on whether that person’s commitments entail certain actions, judgements, or norms. You may think that when I promised to help you move out of your apartment, I promised to help you all day, whereas I may think I only owe you a few hour’s work. We may also be asked to justify our attributions. I can be asked why I am leaving when the moving van is only half full, and I may reply by saying something like “I told you when I agreed to do this that I had another engagement in the afternoon.” Now just as any person’s evaluations of their assertions about the world could be wrong, any person’s evaluations of their assertions

about a commitment or entailment can be wrong. Therefore no matter how a person or group evaluates their assertions about a commitment or entailment, other evaluations are possible. This allows for objectivity in the first sense about norms. The slippage between different attributions of commitment also allows for de re ascriptions of commitment and entitlement. Commitments and entitlements, and hence norms in general, are objects that exist apart from our descriptions of them. Imagine that Susan is committed to helping Stan move. Unknown to Susan, a commitment to helping Stan move is a commitment to assist him in carrying 40 large boxes of books from his fourth floor apartment to a U-Haul truck. Suppose further that Susan has a belief about her commitment to Stan: she thinks it can be discharged in half a day. Under these circumstances we can make a de re ascription to Susan. We can say that she believes, of her commitment to assist Stan in carrying 40 large boxes of books from his fourth floor apartment to his U-Haul, that it will take half a day. Since norms are objective in this sense, their relationships of entailment among each other and to actions and judgements are also objective in this sense. This fulfills the second requirement of objectivity.

If normative judgements are right or wrong apart from anyone saying they are right or wrong, and mathematical statements are normative, then it seems that it should be possible for everyone to be wrong about a mathematical statement, in total contradiction with what was asserted in Chapter 2. The conflict is only apparent, however, because it rests on a confusion of the two ways a statement can be an expression of a norm that I outlined above. The impossibility of consistent error about a mathematical statement arises because consistently making the same error means that one is actually following a

different rule. Dissent is still possible in this situation, however. Dissent in this case means advocating a different rule, not advocating a different application of the same rule. If I lived in a world where everyone said that 12×12 is 145, I could still dissent from this consensus. My dissent would take the form of arguing that " $12 \times 12 = 144$ " will give us a more workable multiplication table. The problem, therefore, with objection that the objectivity of norms conflicts with what was said in Chapter 2 is that it expects the kind of dissent that would come with a hybrid normative statement, when in fact the proper kind of dissent is what would come with a pure normative statement. The objection expects dissent to take the form of the claim that commitment to a rule implies a different sort of judgement. It expects us to say to the community where 12×12 is 145 that, applying the same rule, there result should be 144.

It should also be noted that dissent from the application of mathematical norms is possible if one charges that the rule follower is not being consistent. The impossibility of error is only the impossibility of universal, consistent error, because only then do we say that a different rule is being applied. If I claim that someone has made a one-time error, then I can still maintain that the original rule is being followed. There will be gray areas, of course, places where it is not clear if one rule is being followed inconsistently or if two rules are being employed at different times. However, as long as there are cases where it is clear that universal error regarding one rule becomes the correct use of another rule, the claim in Chapter 2 stands.

Cases like the people who say that 12×12 is 145 can be generated for ethical norms, although nothing like this situation exists for descriptive statements. We can easily

imagine a community that says they practice monogamy, but where everyone openly keeps lovers on the side. These people are actually following a rule, just not the rule we ordinarily call “monogamy.” Dissent in this community will always take the form of advocating a new definition of “monogamy.” On the other hand, nothing like this exists for descriptive statements. Descriptive statements allow for universal and consistent error without change in the meaning of a sentence. If everyone believes “all swans are white,” one can dissent from this consensus without changing the meaning of any of the terms of the proposition. Justification for this dissent only requires one to produce a black swan.

The objectivity of normative statements also allows us to explain why they look like descriptive statements. One might think that acceptance of the claim that mathematical statements are actually expressions of norms would lead to a movement to reform mathematical language. Perhaps the new movement would advocate placing all mathematical statements in the subjunctive. This would be very much in conflict both with Wittgenstein’s intentions for the normative thesis and with what the thesis actually entails. What we are looking at here is a specific instance of a basic tension that runs throughout Wittgenstein’s work: the desire on the one hand to say that philosophic problems have arisen because we have been misled by the grammar of a sentence, and on the other hand to say that natural language is perfect as it is and does not need to be reformed. This tension almost always arises because Wittgenstein finds that a certain class of utterances that appear to be descriptive are really in some way “expressive.” The three culprits Wittgenstein talks most about are religious and moral language, mental language, and our topic, mathematical language. In each case our language appears to

describe a realm of rather obscure objects—God and the afterlife, sensations and expectations, numbers and functions—but really is used by the speaker to express something. Talk of sensations is an expression of what a person is feeling. “There is a pain in my leg,” is really of a piece with, “ouch.” Religious and moral talk is a way of expressing certain mystical feelings and experiences. Mathematics, I have argued, is a way of expressing certain norms regarding objects. But if the language used in all of these situations is deceptive, wouldn’t it be better if we spoke in a purified language? Wittgenstein insists that he is not a reformer: “It is not that a new building has to be erected, or that a new bridge has to be built, but that the geography, *as it now is*, has to be described” (*RFM* V 54). Wittgenstein believes that language is properly formed as it is. If this is right, we ought to be able to find a reason for our use of descriptive language in mathematics. Moreover, we ought to be able to justify descriptive language in such a way that we will be able to talk about mathematical objects without making winking asides to the effect that we don’t really mean what we say.

The reason we are justified in using descriptive language to express norms is that norms exhibit the same sort of objectivity that empirical objects do. Thus Brandom writes, “There are normative facts of the matter as well as nonnormative ones. Acknowledging the distinction between them does not make one of them nonobjective, and it does not commit one to distinguishing the natural sciences as achieving some sort of stronger objectivity (objectivity on steroids)” (1997, 203). The solution to the problem of the use of descriptive language in mathematics is to regard mathematical statements as describing normative facts. Numbers, on this account, are evaluative predicates applied to

collections of objects.⁴⁸ The collections are not judged “good” or “bad,” however. They are judged to be collections that we must treat in a certain way. The actions required of us by these evaluations are simply the sorts of actions Wittgenstein points to in his examples. If I put two pairs of apples on the table and find that there are only three apples in total, I am committed to the belief that one of the apples disappeared. Mathematical statements, as expressions of norms, describe the convolutions of commitments and entitlements that the use of these predicates brings. Because statements about commitments and entitlements enjoy the same kind of objectivity that empirical statements do, the statements of mathematics will be as robustly descriptive as any empirical statement. Now it is true that the evaluative predicates are applied to collections after we have interacted with them in a very specific way—counting them. In this sense the predicates applied to collections are evaluative in the same way a predicate like “poisonous” is evaluative. “Poisonous” too is also applied to objects on the basis of interaction with them and demands that the objects in the predicate’s extension be treated a certain way. But in showing that mathematical statements are expressions of norms, we show that mathematical knowledge focuses solely on the commitments and entitlements brought on by the use of the evaluative term and says nothing about the object. Mathematical statements are thus similar to statements about what it means for something to be poisonous, rather than empirical statements about poisonous compounds.

⁴⁸ In what follows, I am only defending the claim that numbers are *evaluative* predicates, not the claim that they are predicates applied to collections of objects. My reply to Frege’s critique of the idea that numbers are predicates applied to objects was given in Chapter 1. (See pp. 52–57.)

Although I have shown that expressions of norms are objective and claimed that we can take the descriptive form of mathematical language at face value, I still have not normally outlined what would ordinarily be called a realist position. In fact, my position seems to be straightforward social constructivism. I have claimed that mathematics describes normative facts, but I have also endorsed Brandom's analysis of normativity, under which normativity is socially instituted. Normative facts arise out of the way individuals evaluate each other in the practice of deontic scorekeeping. For this reason, one still may accuse me of wanting to turn mathematics departments into rhetoric departments. Specifically, one might say that if mathematical facts are socially instituted, some form of relativism follows. Would mathematics be true in societies that do not institute mathematical norms? This concern is bolstered by the fact that at least one interpreter of Wittgenstein's philosophy of mathematics, David Bloor, endorses Wittgenstein's views precisely because he believes they lead to a radical form of relativism.⁴⁹

Bloor is best known for advocating relativist methods in the history of science as a part of what he termed the "strong program" (Bloor 1976; Barnes and Bloor 1982). The goal of the strong program is to identify the causes of knowledge. Its central, and most controversial, tenet is that true and false beliefs are to be treated symmetrically: the same

⁴⁹ I became aware of Bloor's most recent book *Wittgenstein, Rules and Institutions* (1997) too late to include a detailed discussion of it in this dissertation. In that book, however, he avoids talk of relativism, instead talking of "linguistic idealism." Whether or not this represents a change in position, it certainly illustrates his knack for openly advocating ideas that others go to great lengths to distance themselves from.

sorts of causes must be used to explain both. The truth of a belief is not a factor that can be used in explaining its acceptance. In practice, the causes of belief identified are features of the social milieu in which the belief develops, and the causes that are ignored because they entail invoking the truth of the belief are features of the objects in the world the belief purports to be about. Underlying this program is the conviction that knowledge is always knowledge relative to some local, socially imposed standard of justification. The result is a strong form of relativism. Knowledge only exists relative to a social context, and to explain the content of knowledge we must look at the social context. For the strong program, mathematical knowledge is a kind of test case. Because mathematical knowledge appears to be universal, it belies the conviction that knowledge only exists relative to a local standard of justification. Because mathematics is so abstract, it is hard to find causes for it within social structures—there is no obvious connection between the Pythagorean theorem and any aspect of ancient Greek society. The challenge posed by mathematical knowledge led Bloor to try to develop a sociology of mathematics, and to the work of Ludwig Wittgenstein. Wittgenstein, according to Bloor, has shown us that mathematics is a product of social agreement, just as the sociology of science shows that the empirical sciences are the result of social forces coming to a balance. A key example of this is the case of the pupil asked to count by twos. Under Bloor's interpretation Wittgenstein has shown that the correct way to extend the series cannot be found in any Platonic realm of number, or in the empirical world. This means that the correct extension is "entirely dependent on instinct, training and convention" (1983, 87). We extend the series the way we do because of our make-up and the way that make-up is

shaped by society. The fact that we all agree on how to extend the series is a product of our shared society and common set of instincts. For Bloor, the normativity of mathematics is another guise of its social constitution. Saying that mathematics is normative is another way of saying that it derives its compelling force from “being accepted and used by a group of people” (ibid., 92). This leaves Bloor free to practice the sociology of mathematical knowledge the way he sees fit. He can, for instance, apply Mary Douglas’s ideas about the way societies respond to outside threats to the counterexamples provided to early proofs of Euler’s theorem (Bloor 1978).

Bloor’s brand of relativism has attracted genuine outrage. To many, the idea that scientific knowledge is socially constituted is not just wrong, it is offensive, a sign of social collapse and anarchy in the academy. If the account I have advocated were grouped together with Bloor’s, I might be in for some heavy criticism. Fortunately, there is a key difference between the conception of normativity I have adopted and the conception Bloor operates with, a difference that keeps my account from falling into Bloor’s relativistic excesses. For Bloor, norms are simply patterns of behavior adopted by a community. In Brandom’s terms, Bloor’s normativity consists solely of the commitments and entitlements that the majority of its members attribute to themselves. This means that Bloor’s norms are not objective in either of the two senses we have been dealing with. There is no possibility of everyone being wrong about their normative commitments, because their commitments simply are what most people acknowledge them to be. There is also no possibility of *de re* ascription of commitments. If no one acknowledges that two commitments are equivalent, then one cannot say that anyone’s

beliefs about the one commitment are beliefs de re about the other, because for Bloor, if the equivalence is unacknowledged, it doesn't exist. Because Brandom's account gives norms a greater objectivity than Bloor's does, Brandom's account allows us to say that mathematical facts are true even for communities where people do not hold mathematical beliefs. Such facts amount to unacknowledged commitments and entitlements. Brandom's account also allows for what Bloor would consider symmetry breaking in the explanation of the cause of knowledge, in that the same facts that justify a historian's belief in a mathematical fact may also be used as explanations for other's belief in that fact. In providing an explanation as to why a person has followed an ethical norm, it is standard to appeal to ethical facts that person was cognizant of. Similarly, in explaining a mathematical belief, one can appeal to mathematical facts they were cognizant of. Thus, for instance, if we wanted to explain Euclid's belief in the Pythagorean theorem, we can appeal to the proof that was given to him when he first learned the theorem from his teachers, which would probably be some version of the proof he set down as proposition 47 of Book 1 of the *Elements*.

The Universality of Mathematical Norms

There is one final worry one might have about the objectivity of mathematics under Wittgenstein's approach. We have shown that expressions of norms can be objective, in that mathematical statements are true or false apart from anyone's say so, and can be used to refer to extralinguistic entities. We have also justified the use of descriptive language in mathematics and what Bloor would call asymmetric explanations in the history of

mathematics. However mathematics still does not have the universality typically attributed to it. One might fear that a relativist or other postmodern bogeyman could use the normative approach to mathematics to paint a picture of higher mathematics as a uniquely European or white male product. If mathematics is an expression of norms, then wouldn't it really be an expression of the norms of the dominant society? To counter this final worry, I will argue that mathematical language is compatible with any language that talks about objects. Any time you can talk about things, you can count those things. This fact does not rule out the possibility of explaining particular mathematical theories as the product of a particular culture. However, given that all human languages seem to allow for talk of objects, it does allow us to assert that mathematics as a whole is cross cultural.

The claim that mathematical language is compatible with any language does not contradict my earlier claim about the objectivity of expressions of norms. The possibility of dissent allows us to argue that any given statement does not actually express a norm we ought to follow. It also allows us to say that any given action is not a correct application of a norm, unless the norm is being consistently applied in this fashion. None of this conflicts with the possibility that the language we use to frame this debate is universal. Indeed, if I am correct, robust dissent from mathematics as a whole should be incomprehensible. One cannot advocate the use of a language that forbids mathematical talk, unless it is highly impoverished.

Mathematical norms bear a special relationship to norms involving the object concept. The intuition behind this claim can be found in Wittgenstein's remarks about attempts to collect objects that go wrong. If we put two pairs of apples on the table and

find that we only have three apples (as in *RFM* I §37), we would say that one disappeared, or somehow merged with the others, like they were drops of water. In this example, as in all of Wittgenstein's examples, violations of mathematical norms seem to lead to violations of object norms. Objects cease to exist or merge into each other. The intuitive link between object norms and mathematical norms is also in play in Karen Wynn's experiments with infant recognition of small collections. As you recall, she tested infant's abilities to recognize the size of collections by placing collections of objects behind a screen and monitoring infant reaction when the screen was lifted to reveal a different number of objects. Here again the violation of the arithmetic norm is achieved through violation of object norms: to create the anomalous cases, Wynn had to surreptitiously make objects disappear. Indeed, some objected that her experiments did not test for arithmetic cognition at all, but for knowledge of object permanence.

I would like to make the intuition that number norms and object norms are linked more rigorous. I will argue that any language compatible with the use of singular terms must also be compatible with basic mathematical language. Specifically, it must allow the members of collections to be numbered. My argument will be an extension of Brandom's argument that any language capable of expressing either the conditional or negation must be capable of describing objects. So my actual conclusion will be even stronger than what I have stated; really, any language capable of expressing either the conditional or negation must be compatible with mathematical language. Note that I am not claiming that the language will in fact have number words, any more than Brandom is saying that every language must in describe objects. Both arguments are merely about what may be

introduced to a language without contradiction. Brandom's argument works by showing that if a language is compatible with basic logical terms and it has subsentential structures, those structures must include singular terms. My extension will show that if such a language is able to express basic features about the way its singular terms behave, it must also be able to express basic arithmetic facts about collections of objects. It must be able to distinguish equivalence classes of collections based on their numerosity and place an ordering relation on those equivalence classes.⁵⁰

Let's begin with Brandom's deduction of the necessity of objects. Brandom's approach to language is centered around the fact that linguistic acts are things that can be evaluated normatively. As a result, he takes the sentence to be the primary unit of language, because sentences are the smallest unit that can constitute a move in a language game, and are thus the main center for normative evaluation. Among sentences, assertions are primary because they can both serve as reasons and require reasons to back them up: They can both justify a move in a language game and be an invitation to a further move. But although assertions are the beginning of Brandom's analysis of language, they are not the end. If we are going to be able to predict the meaning of novel sentences based on known sentences, we must be able to decompose sentences into

⁵⁰ Gödel's results will not be an issue here because no claim to completeness is being made. I am not attempting to derive an axiomatic system that has all true theorems as its consequence. In fact, an important theme in Brandom's work that I am accepting is that there can be no complete statement of the norms implicit in language. Viewed this way, the incompleteness theorems merely reinforce a fact that we already knew. Linguistic norms cannot be made fully explicit, so it should come as no surprise that a subset of these norms, the norms surrounding singular terms, cannot be explicitly stated in an axiomatic system.

smaller structures and rearrange these elements. The basic method for such decomposition is substitution. Substitution gives us three basic subsentential structures, the part substituted for, the part substituted in, and the frame that stays the same throughout this process. According to Brandom, the frame is a derived category, something which arises from the other two structures. As a result, changing the frame of a sentence is not considered substitution in the proper sense. It is merely replacement. In ordinary languages, the expressions substituted for and substituted in are singular terms, and the frames are predicates. Now the significant feature of assertions was that they serve as reasons, most often as reasons for other assertions. In other words, assertions stand in inferential relations to one another—one assertion stands in an inferential relationship to another if it can serve as a reason for it. We can use these inferential relationships between sentences to define a similar structure of commitments regarding singular terms. If we can substitute one singular term for another in any sentence and maintain the validity of all the inferential relationships of that sentence, we have an important commitment about the relative inferential strengths of the two singular terms. For instance, the singular term “the inventor of bifocals” can be substituted in for “Ben Franklin” in the sentence “Ben Franklin was the first postmaster general of the United States” without altering the inferential relationships of that sentence. Therefore we are in an important way committed to the equivalence of the singular terms “Ben Franklin” and “the inventor of bifocals.” Brandom dubs this commitment a Simple Material Substitution-Inferential Commitment, or SMSIC. SMSICs involving singular terms have the interesting feature of being symmetric. If a is substitutable for b saving the validity of

any inferences, then b must be substitutable for a . We can therefore define equivalence classes of terms that are substitutable for each other. These equivalence classes mark out objects—an object is the thing in the world that corresponds to such an equivalence class. Predicates do not always have the sort of symmetry found in substituting singular terms. One can replace “is a mammal” with “is a dog” in a sentence and all of the old inferences involving that sentence will still be true. For instance, “Fido is a mammal” implies “Fido is warm blooded,” and “Fido is a dog” also implies “Fido is warm blooded.” However, replacing “is a dog” with “is a mammal” invalidates many inferences—for instance “is man’s best friend” no longer follows.

One important result of this way of looking at singular terms is that no object can be named by only one singular term. The crucial feature of being an object is that one can reidentify it later. From the perspective of language, this fact becomes the requirement that an object must be named by multiple definite descriptions or proper names. In order for the man I meet on the street on Tuesday to count as an individual, I will have to be able to meet him again, in which case, he will fall under another description, such as “the man I met on the street on Wednesday.” In Brandom’s theory of language, this means that any singular term must be involved in at least one SMSIC with another singular term. “The man I met on the street on Tuesday” and “the man I met on the street on Wednesday” are related by a SMSIC and form two elements of the equivalence class of singular terms that corresponds to that man. A parallel result, which I do not believe Brandom mentions, is that every singular term must also *not* be inferentially equivalent to at least one other singular term. The alternative would be to say that there is some

singular term that is equivalent to all other singular terms. But since inferential equivalence is a transitive relationship, this would mean that every singular term has the same inferential significance. If this consequence does not seem immediately absurd, remember that we are taking the sentence as the primary unit of language, and only concern ourselves with subsentential structures in order to project the meaning of future sentences based on the sentences we know. But if every singular term is substitutionally equivalent to every other singular term, then they will be no good for predicting the meaning of novel sentences. We will have failed to decompose sentences in a useful fashion. Therefore any singular term must be involved in at least one positive and one negative SMSIC.

Brandom shows that all languages able to express conditionals and negation must be compatible with the existence of singular terms by showing that, if such a language has any subsentential structures, they must use substitutables involved in only symmetric SMSICs and frames involved in asymmetric SMSICs. There are four logical possibilities for subsentential structures: either both frame and substitutable are symmetric, the substitutable is symmetric and the frame asymmetric, the frame is symmetric and the substitutable is asymmetric, or both frame and substitutable are asymmetric. The first possibility can be ruled out quickly. Some inferential relations must be asymmetric, otherwise we could never employ inferences where the conclusion is weaker than the premise—inferences from generalities to particulars, for instance. The second possibility is the case that actually obtains—the situation of a language with singular terms. The remaining two possibilities both feature asymmetric substitutables. Brandom's proof

shows that asymmetric substitutables are incompatible with languages capable of expressing the conditional or negation. Let Qa and Qb be two sentences such that one can infer Qb from Qa . If there was another predicate Q' such that one can infer $Q'a$ from $Q'b$, it would be impossible for a SMSIC to govern any of these inferences. No symmetric SMSIC governs them because by hypothesis there are no symmetric SMSICs. But no asymmetric SMSIC could govern them either. An asymmetric SMSIC must either declare a or b to be the inferentially stronger term. But this means it can only license the inference from Qa to Qb , or the inference from $Q'b$ to $Q'a$, but not both. Now if there were an algorithm that could generate a Q' for any predicate and any pair of singular terms, then it would be impossible for any SMSICs to govern the use of any singular term. Brandom offers two such algorithms. First, given Q , let Q' be $Q \rightarrow r$, where r is any sentence. Second, given Q , let Q' be $\sim Q$. Both these algorithms deserve close attention.

Brandom claims that if Qa implies Qb , then $Qb \rightarrow r$ will imply $Qa \rightarrow r$. Brandom explains this by saying: “Because conditionals make inferential commitments explicit as the contents of assertional commitments, inferentially weakening the antecedent of a conditional inferentially strengthens the conditional” (1994, 380).⁵¹ Brandom’s second

⁵¹ The converse is not necessarily true. One might think that this matters, but it doesn’t. Symbolically, Brandom’s argument is represented $Qa \rightarrow Qb \vdash (Qb \rightarrow r) \rightarrow (Qa \rightarrow r)$. One might think he needs to show something stronger, namely, that $\sim(Qb \rightarrow Qa) \cdot \sim(Qb \rightarrow Qa) \vdash (Qb \rightarrow r) \rightarrow (Qa \rightarrow r)$ and $Qa \rightarrow Qb \cdot \sim(Qb \rightarrow Qa) \vdash \sim[(Qa \rightarrow r) \rightarrow (Qb \rightarrow r)]$. The latter simply is not true, as is shown by a model with a two element universe of discourse $\{\mapsto, \uparrow\}$, where the extension of Q is \uparrow and r is assigned the value T. Brandom would need the latter entailment if his initial claim went like this: Let Qa and Qb be two sentences such that one can infer Qb from Qa , *but not vice versa*. If there was another predicate Q' such that one can infer $Q'a$ from $Q'b$, *but not vice versa*, it would be impossible for a SMSIC to govern any of these inferences. Brandom does not need the

algorithm says that given Q , let Q' be $\sim Q$. Obviously, the stronger any proposition Qa , the weaker $\sim Qa$. Therefore if Qa implies Qb and Qb does not imply Qa , $\sim Qa$ will not imply $\sim Qb$, but $\sim Qb$ will imply $\sim Qa$. Given either of these algorithms, Brandom's conclusion follows. If the implicit practices of a language allow for the introduction of either the conditional or negation, any subsentential structure that language may have must involve singular terms and predicates. This does not mean that every language must have such structures explicitly. One need not ever use logical vocabulary or singular terms. However, if a language is to be capable of accepting one it must be capable of accepting the other.

Brandom's proof can be extended to show not only that languages capable of accepting basic logical vocabulary must have singular terms, but that the objects named by these singular terms must be numerable, that is, it must be possible to divide collections into equivalence classes that contain at least two members and that we can define a successor relationship on. What would it be for objects not to behave this way? They would be objects that one could not make into stable collections. The objects would slip away or blend into each other, before one had a chance to say anything definite about the way they are grouped. For a language only to name objects of this kind, it would have to parse the world in such a way that it could not reidentify the number of a collection. Since all we know about objects in any possible language right now is that they are represented by singular terms, all we know about collections of objects is that they can be

“but not vice versa” clauses, however. They only serve to rule out symmetric SMSICs, which have already been ruled out by hypothesis.

represented by collections of singular terms. Our goal, then, is to show that all collections of singular terms can be divided into equivalence classes based on the number of objects named by the collection of singular terms. This is the first adequacy requirement on our venture. The second is that we do not beg the question by assuming some notion of number. Only abilities associated with the use of singular terms and other vocabulary Brandom considers logical or explicatory may be used in our proof. Specifically, let's give ourselves the following logical tools:

1. Singular terms divided into equivalence classes representing objects: $a_1, a_2, a_3 \dots$; $b_1, b_2, b_3 \dots$ etc. Variables ranging over singular terms will be in italics— x, y, z etc.
2. The two place predicate “__ is distinct from __,” which is true of singular terms x and y iff they belong to distinct equivalence classes.
3. The standard sentential connectives $\&$, \vee , \sim , \rightarrow and \leftrightarrow .
4. The ability to create collections of singular terms. Minimally, a language could do this simply by listing singular terms.⁵² As a shorthand, we will refer to collections of singular terms using boldface letters: **a**, **b**, etc. Variables ranging over collections of singular terms will be in boldface italics: ***x***, ***y***, ***z*** etc.
5. The ability to consistently substitute collections of singular terms into sentence frames with multiple variables, and perhaps multiple occurrences of the same variable.

Our goal is to divide collections of singular terms into equivalence classes that we would intuitively call equinumerous. Traditionally, the sizes of collections are picked out via the notion of one-to-one correspondence. (This is what Frege does in the *Grundlagen* [1884/1968, §68].) The notion of one-to-one correspondence is also a crucial point where

⁵² We do not need to worry about comprehension principles, because we will never be generating sets of singular terms according to rules. We are only concerned with finite collections of arbitrary terms. Nor do we need to worry about what sortal the objects are gathered under. We can just assume that they are gathered under some sortal.

the object concept intersects with the number concept. Intuitively, what allows us to put two collections in one-to-one correspondence is that the objects in the set maintain their distinctness from one another—they do not merge or disappear. In dealing with singular terms, we need to find a linguistic equivalent to the physical process of putting objects in one-to-one correspondence. I think the needed linguistic equivalent is a set of frames, which I will call the frames of *type D*, and which is defined by the following multistage recursive process:

- (df. 1) 1. All frames of the form $\lceil \sim(x \text{ is distinct from } y) \rceil$ are of type D. Frames allowed into type D by this rule are *stage 1 frames*.
2. If ϕ is a frame of the form named in (i), then $\sim\phi$ is a frame of type D. Frames allowed into type D by this rule are *stage 2 frames*.
3. If ψ is a frame of type D not formed at stage 1, the χ is also a frame of type D iff χ is the result of expanding ψ via conjunction of frames from stage 2, where
- a. all of the new conjuncts contain one singular term from ψ and one not from ψ .
 - b. The new singular term must be the same in each conjunct.
 - c. The old singular term must be different in each conjunct.
 - d. One conjunct must be introduced for each singular term in ψ .⁵³
4. These are all the frames of type D.

These rules will generate frames that assert the existence of more distinct objects at each stage of recursion. Stage two gives us frames of the form, $\lceil \sim\sim(x \text{ is distinct from } y) \rceil$ or simply $\lceil (x \text{ is distinct from } y) \rceil$. Stage three will have frames of the form $\lceil (((x \text{ is distinct from } y) \ \& \ (x \text{ is distinct from } z)) \ \& \ (y \text{ is distinct from } z)) \rceil$. Stage four will have frames of

⁵³ For convenience we can say that if ψ is formed at stage n , then χ is formed at stage $n + 1$, although this is not strictly necessary.

the form $\lceil (((x \text{ is distinct from } y) \ \& \ (x \text{ is distinct from } z)) \ \& \ (y \text{ is distinct from } z)) \ \& \ (((w \text{ is distinct from } x) \ \& \ (w \text{ is distinct from } y)) \ \& \ (w \text{ is distinct from } z)) \rceil$, etc. It is important to note that the numbering of the stages is strictly a convenience. All we need for the purposes of our definition is names for the first two stages that will allow us to specify how conjunctions are formed in the other stages. If we relied on the numbering in any more substantial fashion, we would of course be at risk of begging the question.

Next, equivalence classes need to be generated based on how collections of singular terms can be substituted into such frames. Following Brandom, we will not be concerned with what we can substitute into a sentence frame *salva veritate*. Rather, we will develop an inferential semantics, in which what is important about a sentence is not its truth, but what one can infer from it. Therefore we will partition our collections of singular terms into classes based on whether substituting them into frames of the above sort preserves inferential significance. In doing this, however, we will rely on the fact that in classical logic falsehoods have a unique inferential significance: anything follows from them. I will therefore say that that a collection of singular terms can be *properly substituted* into a sentence frame if it can be substituted in without yielding a sentence that has the inferential significance of a falsehood. Given this bit of terminology, let's define the relationship C as follows:

(df. 2) **aCb** iff

1. **a** and **b** can be properly substituted into some frames of type D
2. **a** and **b** can be substituted into all and only the same frames of type D

C is not an equivalence relation because it is not reflexive. If **a** cannot be substituted into any frame of type D, then $\sim\mathbf{aCa}$. But this will not be a problem.

Not all collections of singular terms fall into a relationship C . Any collection of more than two terms where some terms name the same object is ruled out of consideration by the specification that the collection must be *properly* substitutable into a frame of type D . To deal with these terms we need to define another relation C' .

(df. 3) $\mathbf{aC'b}$ iff

1. One or both of the pair \mathbf{a} and \mathbf{b} cannot be properly substituted into any frame of type D .
2. Removing some singular terms from the collections that cannot be substituted into any frame of type D yields collections that can be properly substituted into some frames of type D
3. Removing the smallest number of singular terms from the collections that cannot be properly substituted into a frame of type D necessary to yield collections that can be properly substituted into a frame of type D yields a pair of collections that stand in relationship C to each other.

Like C , C' fails to be an equivalence relation because it is not reflexive. If \mathbf{a} is properly substitutable into a frame of type D , then $\sim\mathbf{aC'a}$, by the first criterion.

The classification we need is one based on an equivalence relation that unites C and C' . Lets call this relation C'' .

(df. 4) $\mathbf{aC''b}$ iff either \mathbf{aCb} or $\mathbf{aC'b}$

C'' is an equivalence relation. To see that $\mathbf{aC''a}$, we need to consider two cases. Either \mathbf{a} can be substituted into a frame of type D , or it can't. If it can, then $\mathbf{aC'a}$ because \mathbf{a} will properly fit into all and only the same frames of type D as itself. If \mathbf{a} cannot properly fit into a frame of type D , then it must be possible to remove elements from it to make it properly fit. There are two ways \mathbf{a} might not be properly substitutable into a frame of form D . The collection "Tom, Dick, Harry" cannot be properly substituted into the frame $\lceil \sim(x \text{ is distinct from } y) \rceil$ because it has more members than the frame has variables. On the

other hand, the collection “Wittgenstein, Russell, the author of the *Why I Am Not a Christian*” cannot be properly substituted into the frame $\lceil(((x \text{ is distinct from } y) \ \& \ (x \text{ is distinct from } z)) \ \& \ (y \text{ is distinct from } z))\rceil$ because the substitution yields a falsehood. Now we know that **a** must be substitutable into some frame in the way that “Wittgenstein, Russell, the author of the *Why I Am Not a Christian*” is substitutable into the frame $\lceil(((x \text{ is distinct from } y) \ \& \ (x \text{ is distinct from } z)) \ \& \ (y \text{ is distinct from } z))\rceil$, because we can generate frames of arbitrarily large size, and therefore can find one large enough to fit the number of members of **a**. The problem must then be that **a** yields a sentence with the inferential significance of a falsehood when substituted into the frames that it can be substituted into. This can only occur if two members of **a** are not distinct. But if this is the case, then removing the duplicate names must eventually yield a sentence that can be properly substituted into a frame of type D. The worst case scenario is that all of the elements of **a** name the same object, in which case I could remove all the elements save two and then substitute the resulting collection properly into the sentence frame that forms the root of our recursive tree, $\lceil\sim(x \text{ is distinct from } y)\rceil$. Therefore it must be possible to reduce the members of **a** until it fits in some frame of type D. But once the requisite elements are removed, the reduced **a** will fit into all and only the same frames as itself. Therefore $\mathbf{aC''a}$. The symmetry and transitivity of C'' can be proven similarly.

The crucial step is now to show that any collection of singular terms **a** must fall into a class C'' with some collection besides itself. We can divide the problem up into cases. Either **a** is properly substitutable into some frame of type D, or it isn't. Suppose it is. I say that I can find some singular term, not in **a**, which is distinct from sufficiently many other

objects so that it and those other objects can form a collection that stands in relationship C to **a**. The thing to bear in mind here is that this new singular term need not successfully refer in order to be distinct from other objects. A unicorn is not a centaur, even if neither exist. To create the needed singular term, we can therefore begin by inventing a new name, call it x . Given access to the negation sign, and the ability to substitute x into the sentence frames that the terms in **a** occur in, we can stipulate that whatever is true of any member of **a** is not true of x . To round out the referential purport of x , we can introduce another singular term y to stand in a SMSIC with x . We do not need to know anything about this term save that were there to be a y , it would be the same as x . (This is like saying, were there a Hamlet, he would be the son of Gertrude.) But given the existence of x , I can form a collection with x and all of the elements of **a** save one. This collection will be substitutable into all of the frames of type D **a** is. Therefore there is a collection that stands in relationship C to **a**, and hence in C" to **a**, after all.

The remaining case is easy. Suppose on the other hand that **a** cannot be properly substituted into a frame of type D. As we demonstrated when we showed C" to be an equivalence relation, it must be possible to reduce **a** into a collection **a_{reduced}** that is substitutable into a frame of type D. But, by what we have just shown, we can easily find a collection that stands in relationship C to **a_{reduced}**. By the definition of C", this collection must stand in relationship C" to **a**. Therefore, I conclude that for any collection of singular terms **a** there must be a collection **b** that is distinct from **a** and that stands in relationship C" to **a**.

We have therefore partitioned all of the collections of singular terms of a language

into equivalence classes. From here, it is no problem to introduce an ordering relationship on the set of equivalence classes defined by C'' . Let the script letters \mathfrak{a} and \mathfrak{b} refer to equivalence classes based on the relationship C'' . We can define the relationship O as holding iff adding a singular term to some member of \mathfrak{a} , \mathfrak{a} , which does not refer to any object named in \mathfrak{a} , yields a collection which stands in relationship C to the members of \mathfrak{b} .

The main problems I need to be concerned with are possible objections to my proof that collections of singular terms can be divided into equivalence classes by C'' . One possible source of concern is that I have not really drawn on any of the important inferential features of language that Brandom's proof brings out. To the extent that this is true, I embrace it. I think that the link between the number concept and the object concept is not an artifact of Brandom's system. I am only working in the context of Brandom's theory of language because I think it is the right one to bring out this relationship. I do, however, lean on Brandom's ideas in at least one key place. The central element in my proof is the predicate “_ is distinct from _”, and I have defined this explicitly in terms of Brandom's understanding of objects as corresponding to equivalence classes among singular terms.

A more concrete worry has to do with my use of the word “stipulate” in finding a \mathfrak{b} that stands in relationship C to a collection \mathfrak{a} that can be substituted into a frame of type D . I created the necessary \mathfrak{b} by stipulating a new singular term that purports to refer to an object not named by \mathfrak{a} . But Brandom, it will be objected, says that anyone committed to an expressive theory of logic cannot use stipulative methods. If we are attempting to bring out what is implicit in existing language, it does us no good to simply introduce

new languages. However, I am not stipulating in the way Brandom finds objectionable. I am not making up an ideal language. When I say that we can stipulate an x that has none of the properties of the members of a set \mathbf{a} , I am merely showing what any language user can do. I am working within the existing language, not making up a new one.

References

- Abbagnano, Nicola. 1967. "Psychologism." In *Encyclopedia of Philosophy*. Edited by Paul Edwards. Translated by Nino Lagiulli. New York: McMillan.
- Ambrose, Alice, and Morris Lazerowitz. 1948. *Fundamentals of Symbolic Logic*. New York: Rinehart & Co.
- Anderson, S. W., A. R. Damasio, and H. Damasio. 1990. "Troubled Letters but Not Numbers: Domain Specific Cognitive Impairments Following Focal Damage in Frontal Cortex." *Brain* 113: 749–66.
- Antell, S., and D. Keating. 1983. "Perception of Numerical Invariance in Neonates." *Child Development* 54: 694–701.
- Appolonio, I., L. Rueckert, A. Partiot, I. Litvan, J. Sorenson, D. Le Bihan, and J. Grafman. 1994. "Functional Magnetic Resonance Imaging (F-MRI) of Calculation Ability in Normal Volunteers." *Neurology* 44 (Supp. 2): 262.
- Ashcraft, Mark. 1992. "Cognitive Arithmetic: A Review of Data and Theory." *Cognition* 44: 75–106.
- Ashcraft, Mark H., and J. Battaglia. 1978. "Cognitive Arithmetic: Evidence for Retrieval and Decision Processes in Mental Addition." *Journal of Experimental Psychology: Human Learning and Memory* 4: 527–38.
- Baker, G., and P. M. S. Hacker. 1980 and 1985. *Wittgenstein: Understanding and Meaning. An Analytical Commentary on the Philosophical Investigations*. 2 vols. Chicago: University of Chicago Press.
- . 1984. *Scepticism, Rules and Language*. Oxford: Basil Blackwell.
- Barnes, Barry, and David Bloor. 1982. "Relativism, Rationalism, and the Sociology of Knowledge." In *Rationality and Relativism*. Edited by M. Hollis and S. Lukes. Cambridge: MIT Press.

- Baumann. 1868. *Die Lehren von Zeit, Raum und Mathematik*, vol. 1. Berlin.
- Beck, Lewis White. 1955/1965. "Can Kant's Synthetic Judgements Be Made Analytic?" *Kant-Studien* 47. Reprinted in *Studies in the Philosophy of Kant*. Indianapolis: Bobbs-Merrill.
- Benacerraf, Paul. 1973. "Mathematical Truth." *Journal of Philosophy* 70: 661–79.
- Benfield, David. 1974. "The A Priori–A Posteriori Distinction." *Philosophy and Phenomenological Research* 34: 151–66.
- Berger, H. 1926. "Über Rechenstörungen bei Herderkrankungen des Grosshirns." *Archive für Psychiatrie und Nervenkrankheiten* 78: 238–63.
- Biersack, Aletta. 1982. "The Logic of Misplaced Concreteness: Paiela Body Counting and the Nature of the Primitive Mind." *American Anthropologist* 84: 811–29.
- Blackburn, Simon. 1984. "The Individual Strikes Back." *Synthese* 58: 281–301.
- Bleek, Dorothea. 1928/1978. *Naron: A Bushman Tribe of the Central Kalihari*. Cambridge: Cambridge University Press. Reprinted under the same title. New York: AMS Press.
- Bleek, Dorothea. 1937. "Grammatical Notes and Texts in the |auni Language." In *Bushmen of the Southern Kalihari*. Edited by J. D. Rheinallt Jones and C. M. Doke. Johannesburg: University of the Witwatersrand Press.
- Bloor, David. 1976. *Knowledge and Social Imagery*. London: Routledge.
- . 1978. "Polyhedra and the Abominations of Leviticus." *British Journal for the Philosophy of Science* 11: 245–72.
- . 1983. *Wittgenstein: A Social Theory of Knowledge*. New York: Columbia University Press.
- . 1997. *Wittgenstein, Rules, and Institutions*. New York: Routledge.

- Boethius. c.510/1936. *The Theological Tractates*. Translated by H. F. Stewart and E. K. Rand. Cambridge: Loeb Classical Library.
- BonJour, Lawrence. 1998. *In Defense of Pure Reason*. Cambridge: Cambridge University Press.
- Boolos, George. 1971. "The Iterative Conception of Set." *Journal of Philosophy* 68: 215–31.
- Boyd, Richard. 1984. "The Current Status of Scientific Realism." In *Scientific Realism*. Edited by J. Leplin. Berkeley and Los Angeles: University of California Press.
- Brandom, Robert. 1994. *Making It Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge: Harvard University Press.
- . 1997. "Precis of *Making It Explicit* and Responses to Critics." *Philosophy and Phenomenological Research* 57: 153–157, 189–204.
- Broad, C. D. 1936. "Are There Synthetic A Priori Truths?" *Proceedings of the Aristotelian Society* Supp. Vol. XV.
- Cambell, J. I. D., and D. J. Graham. 1985. "Mental Multiplication Skill: Structure, Process and Acquisition." *Canadian Journal of Psychology* 35: 338–66.
- Caramazza, Alfonso, and Michael McCloskey. 1988. "The Case for Single-Patient Studies." *Cognitive Neuropsychology* 5: 517–28.
- Carnap, Rudolf. 1950/1988. "Empiricism, Semantics and Ontology." *Revue Internationale de Philosophie* 4: 20–40. Reprinted as an appendix to *Meaning and Necessity*. Chicago: University of Chicago Press.
- Carroll, Lewis. 1895. "What the Tortoise Said to Achilles." *Mind* 4: 278–80.
- Cartwright, Nancy. 1980/1983. "Truth Doesn't Explain Much." *American Philosophical Quarterly* 17. Reprinted in *How the Laws of Physics Lie*. Oxford: Oxford University Press.

- Casullo, Albert. 1977. "The Definition of A Priori Knowledge." *Philosophy and Phenomenological Research* 38: 220–24.
- Chomsky, Noam. 1980. *Rules and Representations*. Oxford: Basil Blackwell.
- Church, Russell M., and Warren Meck. 1984. "The Numerical Attribute of Stimuli." In *Animal Cognition*. Edited by H. Roitblat, T. Bever, and H. Terrace. Hillsdale N.J.: Earlbaum.
- Cipolotti, Lisa. 1995. "Multiple Routes for Reading Words, Why Not Numbers? Evidence from a Case of Arabic Numeral Dyscalculia." *Cognitive Neuropsychology*.
- Cipolotti, Lisa, Brian Butterworth, and Elizabeth Warrington. 1994. "From 'One Thousand Nine Hundred and Forty-Five' to 1000,945." *Neuropsychologia* 4: 503–9.
- Cipolotti, Lisa, Elizabeth Warrington, and Brian Butterworth. 1995. "Selective Impairment in Manipulating Arabic Numerals." *Cortex* 31: 73–86.
- Clark, Andy. 1993. "Minimal Rationalism." *Mind* 102: 587–611.
- Cohen, Laurent, and Stanislas Dehaene. 1996. "Cerebral Networks for Number Processing: Evidence from a Case of Posterior Callosal Lesion." *NeuroCase* 2: 155–74.
- Damerow, Peter. 1988. "Individual Development and Cultural Evolution of Arithmetical Thinking." In *Ontogeny, Phylogeny and Historical Development*. Edited by Sidney Strauss. Norwood, N.J.: Ablex Publishing Corp.
- Dehaene, Stanislas. 1992. "Varieties of Numerical Abilities." *Cognition* 44: 1–42.
- Dehaene, Stanislas, Emmanuel Dupoux, and Jacques Mehler. 1990. "Is Numerical Comparison Digital? Analogical and Symbolic Effects in Two-Digit Number Comparison." *Journal of Experimental Psychology: Human Perception and Performance* 16: 626–41.
- Dehaene, Stanislas, and Jean-Pierre Changeux. 1993. "Development of Elementary Numerical Abilities: A Neuronal Model." *Journal of Cognitive Neuroscience* 5: 390–407.

Dehaene, Stanislas, and Laurent Cohen. 1991. "Two Mental Calculation Systems: A Case Study of Severe Acalculia with Preserved Approximation." *Neuropsychologia* 29: 1045–74.

———. 1994. "Dissociable Mechanisms of Subitizing and Counting: Neuropsychological Evidence from Simultanagnosia Patients." *Journal of Experimental Psychology: Human Perception and Performance* 20: 958–75.

———. 1995. "Towards an Anatomical and Functional Model of Number Processing." *Mathematical Cognition* 1: 83–120.

———. 1997. "Cerebral Pathways for Calculation: Double Dissociation between Rote Verbal and Quantitative Knowledge of Arithmetic." *Cortex* 33: 219–50.

Dehaene, Stanislas, Laurent Cohen, Nathalie Tzourio, Victor Frak, Laurence Raynaud, Jacques Mehler, and Bernard Mazoyer. 1996. "Cerebral Activations during Number Multiplication and Comparison: A PET Study." *Neuropsychologia* 34: 1097–1106.

Déjerine, J. 1892. "Contribution à l'Étude Anatomico-pathologique et Clinique des Différentes Variétés de Cécité Verbale." *Mémoires de la Société de Biologie* 4: 61–90.

Deloche, G., and X. Seron. 1982a. "From One to 1: An Analysis of a Transcoding Process by Means of Neuropsychological Data." *Cognition* 12: 119–49.

———. 1982b. "From Three to 3: A Differential Analysis of Skills in Transcoding Quantities between Patients with Broca's and Wernicke's Aphasia." *Brain* 105: 179–33.

———. 1987. "Numerical Transcoding: A General Production Model." In *Mathematical Disabilities: A Cognitive Neuropsychological Perspective*. Edited by G. Deloche and X. Seron. Hillsdale N.J.: Erlbaum.

Descartes, René. 1647/1973. *Notae in Programma*. Reprinted as "Notes Directed against a Certain Program" in *Philosophical Works of Descartes*, vol. 1. Edited and translated by Elizabeth S. Haldane and G. R. T. Ross. Cambridge: Cambridge University Press.

Dixon, R. M. W. 1980. *Languages of Australia*. Cambridge: Cambridge University Press.

Duhem, Pierre. 1914/1991. *La Théorie Physique: Son Obejet, Sa Structure*. Paris: Marcel Rivière & Cie. Reprinted as *The Aim and Structure of Physical Theory*. Translated by Philip P. Wiener with an introduction by Jules Vuillemin. Princeton: Princeton University Press.

Dummett, Michael. 1959/1978. "Wittgenstein's Philosophy of Mathematics." *Philosophical Review* 68: 324–48. Reprinted in *Truth and Other Enigmas* Harvard University Press.

———. 1973. *Frege: Philosophy of Language*. New York: Harper and Row.

Einstein, Albert. 1934/1982. "On the Method of Theoretical Physics." In *Mein Welbild*. Edited by Carl Seelig. Amsterdam: Querido Verlag. Reprinted in *Ideas and Opinions*, Edited and Translated by Sonja Bargmann New York: Crown Publishers.

Elman, Jeffrey, Elizabeth Bates, Mark Johnson, Annette Karmiloff-Smith, Domenico Parisi, and Kim Plunkett. 1996. *Rethinking Innateness: A Connectionist Perspective on Development*. Cambridge: MIT Press.

Erdmann, Benno. 1892. *Logik*. Halle a. d. S.: Max Niemeyer.

Fernandes, D. M., and R. Church. 1982. "Discrimination of the Number of Sequential Events by Rats." *Animal Learning and Behavior* 10: 171–76.

Field, Hartry. 1989. *Realism, Mathematics, and Modality*. Oxford: Basil Blackwell.

Fine, Arthur. 1984/1986. "Natural Ontological Attitude." In *Scientific Realism*. Edited by J. Leplin. Berkeley and Los Angeles: University of California Press. Reprinted in *The Shaky Game* Chicago: University of Chicago Press.

Flegg, Graham. 1983. *Numbers: Their History and Meaning*. New York: Penguin.

Fodor, Jerry. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: MIT Press.

Fogelin, Robert. 1976. *Wittgenstein*. New York: Routledge.

Frege, Gottlob. 1884/1968. *Gundlagen der Arithmetik*. Breslau: Wilhelm Koenner. Reprinted as *The Foundations of Arithmetic*. Translated by J. L. Austin. Evanston Ill: Northwestern University Press.

———. 1893/1964. *Grundgesetze der Arithmetik*, Band 1. Jena: Verlag Hermann Pohle. Reprinted as *The Basic Laws of Arithmetic*. Translated by M. Furth. Berkeley and Los Angeles: University of California Press.

———. 1894/1977. “Review of E. Husserl Philosophie der Arithmetik.” *Zeitschrift für Philosophie und philosophische Kritik*. 103: 313–32. Reprinted in *Translations from the Philosophical Writings of Gottlob Frege*. Edited and translated by P. Geach and M. Black. Oxford: Blackwell.

Friedman, Michael. 1992. *Kant and the Exact Sciences*. Cambridge: Harvard University Press.

Galison, Peter. 1987. *How Experiments End*. Chicago: University of Chicago Press.

Gallistel, C. R. 1990. *The Organization of Learning*. Cambridge: MIT Press.

Gallistel, C. R., A. Brown, S. Carey, R. Gelman, and F. Keil. 1991. “Lessons from Animal Learning for the Study of Cognitive Development.” In *The Epigenesis of Mind: Essays on Biology and Cognition*. Edited by S. Carey and R. Gelman. Hillsdale, N.J.: Erlbaum.

Galloway, David. 1992. “Wynn on Mathematical Empiricism.” *Mind and Language* 7: 333–58.

Gazzaniga, M. S., and C. E. Smylie. 1984. “Dissociation of Language and Cognition: A Psychological Profile of Two Disconnected Right Hemispheres.” *Brain* 107: 145–53.

Gelman, Rochel. 1990. “First Principles Organize Attention to and Learning about Relevant Data: Number and the Animate-Inanimate Distinction as Examples.” *Cognitive Science* 14: 76–106.

———. 1993. “A Rational-Constructivist Account of Early Learning about Numbers and Objects.” *The Psychology of Learning and Motivation* 6: 61–96.

Gelman, Rochel, and C. R. Gallistel. 1978. *Child's Understanding of Number*. Cambridge: Harvard University Press.

———. 1991. "Subitizing: The Preverbal Counting Process." In *Memories, Thoughts and Emotions: Essays in Honor of George Madler*. Edited by Kessen, Ortong, and Craik. Hillsdale N.J.: Earlbaum.

———. 1992. "Preverbal and Verbal Counting and Computation." *Cognition* 44: 43–74.

Gettier, Edmund. 1963/1986. "Is Knowledge Justified True Belief?" *Analysis*. Reprinted in *Empirical Knowledge*. Edited by P. Moser. Totowa, N.J.: Rowan and Littlefield.

Gibbon, John, and Russell M. Church. 1984. "Sources of Variance in an Information Processing Theory of Timing." In *Animal Cognition*. Edited by H. Roitblat, T. Bever, and H. Terrace. Hillsdale N.J.: Earlbaum.

Gillies, Donald. 1985. "Nature of Mathematical Knowledge" (book review). *Philosophical Quarterly* 35: 104–7.

Goldman, Alvin. 1979. "What is Justified Belief?" In *Justification and Knowledge*. Edited by G. Pappas. Dordrecht: Reidel.

Greenblatt, S. H. 1973. "Alexia without Agraphia or Hemianopsia: Anatomical Analysis of an Autopsied Case." *Brain* 96: 307–16.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

Hart, W. D. 1975. "Innate Ideas and A Priori Knowledge." In *Innate Ideas*. Edited by Steven Stich. Berkeley and Los Angeles: University of California Press.

Hinrichs, J. V., J. L. Berie, and M. K. Mosell. 1982. "Place Information in Multidigit Number Comparison." *Memory and Cognition* 10: 487–95.

Hinrichs, J. V., D. S. Yurko, and J. M. Hu. 1981. "Two-Digit Number Comparison: Use of Place Information." *Journal of Experimental Psychology: Human Perception and Performance* 7: 890–901.

- Hintikka, Jaakko. 1969. "On Kant's Notion of Intuition (Anschauung)." In *Reflections on Kant's Critique of Pure Reason*. Edited by T. Penelhum and J. J. MacIntosh. Belmont, Calif.: Wadsworth.
- Hittmair-Delazer, M., U. Sailer, and T. Benke. 1995. "Impaired Arithmetic Facts but Intact Conceptual Knowledge—A Single Case Study of Dyscalculia." *Cortex* 31: 139–47.
- Hittmair-Delazer, M., C. Semenza, and G. Denes. 1994. "Concepts and Facts in Calculation." *Brain* 117: 715–28.
- Horowitz, Tamara. 1985. "A Priori Truth." *Journal of Philosophy* 82: 225–39.
- Huddleston, Rodney. 1984. *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.
- Hurford, James. 1975. *Linguistic Theory of Numerals*. Cambridge: Cambridge University Press.
- Husserl, Edmund. 1891/1970. *Philosophie der Arithmetik. Psychologische und Logische Untersuchungen*, Erster Band. Halle a. d. S.: C. E. M. Pfeffer (R. Stricker). Reprinted in *Husserlania*, vol. 12. Edited by L. Eley. The Hague: M. Nijhoff.
- . 1900–1/1970. *Logische Untersuchungen*, 2 vols. Halle a. d. S.: M. Neimeyer. Reprinted as *Logical Investigations*. Translated by J. N. Findlay, New York: Routledge.
- Jevons, W. 1871. "The Power of Numerical Discrimination." *Nature* 3: 281–82.
- Johnson, Mark H., and John Morton. 1991. *Biology and Cognitive Development: The Case of Face Recognition*. Oxford: Basil Blackwell.
- Kant, Immanuel. 1781 and 1787/1969. *Kritik der Reinen Vernunft*. Riga: Johann Friedrich Hartknoch. Reprinted as the *Critique of Pure Reason*. Translated by Norman Kemp Smith. New York: St. Martin's Press.
- Kaufman, E., M. Lord, T. Reese, and J. Volkman. 1949. "The Discrimination of Visual Number." *American Journal of Psychology* 62: 498–525.

- Keil, Frank. 1994. "The Birth and Nurturance of Concepts by Domains: The Origins of Concepts of Living Things." In *Mapping the Mind: Domain Specificity in Cognition and Culture*. Edited by L. Hirschfeld and S. Gelman. Cambridge: Cambridge University Press.
- Kitcher, Philip. 1984. *The Nature of Mathematical Knowledge*. Oxford: Oxford University Press.
- . 1988. "Mathematical Naturalism." In *History and Philosophy of Modern Mathematics*. Edited by P. Kitcher and W. Asprey. Minneapolis: University of Minnesota Press.
- . 1992. "The Naturalists Return." *The Philosophical Review* 101: 53–114.
- Kornblith, Hilary. 1980. "Beyond Foundationalism and the Coherence Theory." *Journal of Philosophy* 72: 597–612.
- Kripke, Saul. 1972. *Naming and Necessity*. Cambridge: Harvard University Press.
- . 1982. *Wittgenstein on Rules and Private Language*. Cambridge: Harvard University Press.
- Laties, V. 1972. "The Modification of Drug Effects on Behavior by External Discriminative Stimuli." *Journal of Pharmacology and Experimental Therapeutics* 183: 1–13.
- Leibniz, Gottfried Wilhem. 1703/1949. *Nouveaux Essais sur l'Entendement Humain*. Reprinted as *New Essays Concerning Human Understanding*. Translated by A. G. Langley. LaSalle Ill.: Open Court Publishing.
- Levin, Harvey, and Paul A. Spiers. 1985. "Acalculia." In *Clinical Neuropsychology*. Edited by K. Heilman and E. Vallenstein. Oxford: Oxford University Press.
- Locke, John. 1689/1964. *An Essay Concerning Human Understanding*. Reprinted under the same title. Edited by A. D. Woozley. New York: Penguin.
- Maddy, Penelope. 1985. "Nature of Mathematical Knowledge" (book review).

Philosophy of Science 54: 312–14.

———. 1986. “Mathematical Alchemy.” *British Journal for the Philosophy of Science* 37: 279–314.

———. 1990. *Realism in Mathematics*. Oxford: Oxford University Press.

———. 1991. “Philosophy of Mathematics: Prospects for the 1990s.” *Synthese* 88: 155–64.

Maingard, L. F. 1963. “Comparative Study of Naron, Heitshware and Korana.” *African Studies [Formerly Bantu Studies]* 22.

Mandler, J. M., and B. J. Shebo. 1982. “Subitizing: An Analysis of Its Component Processes.” *Journal of Experimental Psychology: General* 111: 1–22.

Markman, E. M. 1990. “Constraints Children Place on Word Meanings.” *Cognitive Science* 14: 57–77.

Marler, Peter. 1991. “The Instinct to Learn.” In *Epigenesis of Mind: Essays in Biology and Knowledge*. Edited by S. Carey and R. Gelman. Hillsdale N.J.: Earlbaum.

Marshack, Alexander. 1972. *Roots of Civilization: The Cognitive Beginnings of Man's First Art, Symbol, and Notation*. New York: McGraw-Hill.

Matsuzawa, T. 1985. “Use of Numbers by a Chimpanzee.” *Nature* 315: 57–59.

McCloskey, Michael. 1992. “Cognitive Mechanisms in Numerical Processing: Evidence from Acquired Dyscalculia.” *Cognition* 44: 107–57.

McCloskey, Michael, and Alfonso Caramazza. 1987. “Cognitive Mechanisms in Normal and Impaired Number Processing.” In *Mathematical Disabilities: A Cognitive Neuropsychological Perspective*. Edited by G. Deloche and X. Seron. Hillsdale N.J.: Earlbaum.

McGinn, Colin. 1984. *Wittgenstein on Meaning*. Oxford: Blackwell.

Mechner, Francis M. 1958. "Probability Relations within Response Sequences under Ration Reinforcement." *Journal of the Experimental Analysis of Behavior* 1: 109–121.

Mechner, Francis M., and L. Guevrekian. 1962. "Effects of Deprivation upon Counting and Timing in Rats." *Journal of the Experimental Analysis of Behavior* 5: 463–66.

Meck, Warren, and Russell M. Church. 1983. "A Mode Control Model of Counting and Timing Processes." *Journal of Experimental Psychology: Animal Behavior Processes*. 9: 320–34.

Menninger, Karl. 1958. *Number Words and Number Symbols*. Cambridge: MIT Press.

Mill, John Stuart. 1843. *Science of Logic*. New York: Harper and Brothers.

———. 1865/1973. *Examination of Sir William Hamilton's Philosophy*. London. Reprinted under the same title. Edited by J. M. Robinson. Toronto: University of Toronto Press.

Mohanty, J. N. 1974. "Husserl and Frege: A New Look at Their Relationship." *Research in Phenomenology* 4: 51–62.

———. 1980. *Husserl and Frege*. Bloomington, Ind.: Indiana University Press.

Monk, Ray. 1990. *Ludwig Wittgenstein: The Duty of Genius*. New York: Penguin.

Moore, David, J. Beneson, S. Reznick, and M. Peterson. 1987. "Effect of Auditory Numerical Information on Infant's Looking Behavior." *Developmental Psychology* 23: 665–70.

Moyer, R. S., and T. K. Landauer. 1967. "Time Required for Judgements of Numerical Inequality." *Nature* 215: 1519–20.

Narayanan, A. 1992. "Is Connectionism Compatible with Rationalism?" *Connection Science* 4: 271–292.

Natorp, Paul. 1887/1981. "Über objective und subjective Begundung der Erkenntnis."

Philosophische Monatshetfe 23. Reprinted as “On the Objective and Subjective Grounding of Knowledge,” in the *Journal of the British Society for Phenomenology* 12. Translated by D. Kolb.

Nissen, Hans J. 1988. *The Early History of the Ancient Near East 9000–2000 B.C.* Translated by Elizabeth Lutzeier with K. J. Northcott. Chicago: University of Chicago Press.

Nissen, Hans J., Peter Damerow, and Robert K. Englund. 1993. *Archaic Bookkeeping: Early Writing and Techniques of Economic Administration in the Ancient Near East.* Translated by P. Larsen. Chicago: University of Chicago Press.

Notturmo, Mark Amadeus. 1985. *Objectivity, Rationality and the Third Realm.* Dordrecht: Martinus Nijoff Publishers.

Pap, Arthur. 1958. *Semantics and Necessary Truth.* New Haven: Yale University Press.

Parkman, J. M. 1972. “Temporal Aspects of Simple Multiplication and Comparison.” *Journal of Experimental Psychology* 95: 437–44.

Parsons, Charles. 1969/1983. “Kant’s Philosophy of Arithmetic.” In *Philosophy, Science and Method: Essays in Honor of Ernest Nagel.* Edited by S. Morgenbassen, P. Suppes, and M. White. New York: St. Martin’s Press. Reprinted with new postscript in *Mathematics in Philosophy.* Ithaca, N.Y.: Cornell University Press.

———. 1971/1983. “Ontology and Mathematics.” *Philosophical Review* 80: 151–76. Reprinted in *Mathematics in Philosophy* Ithaca, N. Y.: Cornell University Press.

———. 1986. “The Nature of Mathematical Knowledge” (book review). *Philosophical Review* 95: 129–137.

———. 1979–80/1996. “Mathematical Intuition.” *Proceedings of the Aristotelian Society* 80: 145–68. Reprinted in *The Philosophy of Mathematics* Edited by W. D. Hart, Oxford: Oxford University Press.

Pepperberg, Irene M. 1987. “Evidence for Conceptual Quantitative Abilities in the African Grey Parrot: Labeling of Cardinal Sets.” *Ethology* 75: 37–61.

Piaget, Jean. 1952. *The Child's Conception of Number*. New York: Norton.

Poltrock, S. E., and D. R. Schwartz. 1984. "Comparative Judgements of Multidigit Numbers." *Journal of Experimental Psychology: Learning, Memory and Cognition* 10: 32–45.

Quine, W. V. O. 1951/1961. "Two Dogmas of Empiricism." *Philosophical Review* 60: 20–43. Reprinted in *From a Logical Point of View*, 2d ed. Cambridge: Harvard University Press.

———. 1963. *Set Theory and Its Logic*. Cambridge: Harvard University Press.

Resnik, Michael D. 1997. *Mathematics as a Science of Patterns*. Oxford: Oxford University Press.

———. 1980. *Frege and the Philosophy of Mathematics*. Ithaca, N.Y.: Cornell University Press.

Roland, P. E., and L. Friberg. 1985. "Localization of Cortical Areas Activated by Thinking." *Journal of Neuropsychology* 53: 1219–43.

Ryle, Gilbert. 1949. *The Concept of Mind*. New York: Harper and Row.

Sarton, George. 1938. "Prehistoric Arithmetic in Vestonice." *Isis* 23: 462–63.

Saxe, Geoffrey B. 1982. "Developing Forms of Arithmetical Thought Among the Oksamin of Papua New Guinea." *Developmental Psychology* 18: 583–94.

Schmandt-Besserat, Denise. 1992. *Before Writing*, 2 vols. Austin: University of Texas Press.

———. 1996. *How Writing Came About*. Austin: University of Texas Press.

Schwartz, Robert. 1995. "Is Mathematical Competence Innate?" *Philosophy of Science* 62: 227–40.

- Seymour, S. E., P. A. Reuter-Lorenz, and M. S. Gazzaniga. 1994. "The Disconnection Syndrome: The Basic Findings Reaffirmed." *Brain* 117: 105–15.
- Shanker, S. G. 1987. *Wittgenstein and the Turning Point in the Philosophy of Mathematics*. New York: State University of New York Press.
- Shapiro, Stewart. 1983. "Mathematics and Reality." *Philosophy of Science* 50: 523–48.
- Singer, H. D., and A. A. Low. 1933. "Acalculia (Henschen): A Clinical Study." *Archives of Neurology and Psychiatry* 29: 476–98.
- Sokol, Scott M., P. Macaruso, and T. H. Gollan. 1994. "Developmental Dyscalculia and Cognitive Neuropsychology." *Developmental Neuropsychology* 10: 413–41.
- Sokol, Scott M., Michael McCloskey, and Neal Cohen. 1989. "Cognitive Representations of Arithmetic Knowledge: Evidence from Acquired Dyscalculia." In *Cognition in Individual and Social Contexts*. Edited by A. F. Bennett and K. M. McConkey. North-Holland: Elsevier Science Publishers.
- Sokol, Scott M., Michael McCloskey, Neal Cohen, and D. Aliminosa. 1991. "Cognitive Representations and Processes in Arithmetic: Inferences from the Performance of Brain Damaged Patients." *Journal of Experimental Psychology: Learning, Memory, and Cognition* 17: 355–76.
- Spinoza, Baruch. 1677/1992. *Ethica Ordine Geometrico Demonstrata*. Reprinted as *The Ethics*. Translated by Samuel Shirley. Edited by Seymour Feldman. New York: Hackett.
- Starkey, Prentice, and R. Cooper. 1980. "Perceptions of Numbers by Human Infants." *Science* 210: 1033–35.
- Starkey, Prentice, Elizabeth Spelke, and Rochel Gelman. 1983. "Detection of Intermodal Numerical Correspondences by Human Infants." *Science* 222: 179–81.
- . 1990. "Numerical Abstraction by Human Infants." *Cognition* 36: 97–128.
- Steiner, Mark. 1975. *Mathematical Knowledge*. Ithaca, N.Y.: Cornell University Press.

Stich, Steven. 1975. "Introduction." In *Innate Ideas*. Edited by Steven Stich. Berkeley and Los Angeles: University of California Press.

Strauss, M. S., and L. E. Curtis. 1981. "Infant Perception of Numerosity." *Child Development* 52: 1146–52.

Stricker. 1883. *Studiend uber Association der Vorsellungen*. Vienna.

Thompson, Manley. 1972/1992. "Singular Terms and Intuitions in Kant's Epistemology." *Review of Metaphysics* 26: 315–32. Reprinted in *Kant's Philosophy of Mathematics*. Edited by Carl Posy. Dordrecht: Kluwer.

Trick, Lana, and Zenon W. Pylyshyn. 1993. "What Enumeration Studies Can Show Us About Spatial Attention: Evidence for Limited Capacity Preattentive Processing." *Journal of Experimental Psychology: Human Perception and Performance* 19: 331–51.

Van Fraassen, Bas C. 1980. *Scientific Image*. Oxford: Oxford University Press.

Warrington, Elizabeth K. 1982. "The Fractionation of Arithmetical Skills: A Single Case Study." *Quarterly Journal of Experimental Psychology* 31A: 31–51.

Welford, A. T. 1960. "The Measurement of Sensory-Motor Performance: Survey and Reappraisal of Twelve Years' Progress." *Ergonomics* 3: 189–230.

Wigner, Eugene. 1960. "The Unreasonable Effectiveness of Mathematics in the Natural Sciences." *Communications in Pure and Applied Mathematics* 8: 1–14.

Williams, Meredith. 1983/1999. "Wittgenstein on Representation, Privileged Objects, and Private Languages." *Canadian Journal of Philosophy* 13: 57–78. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

———. 1990/1999. "Social Norms and Narrow Content." *Midwest Studies in Philosophy* XV: 425–62. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

———. 1991/1999. "Blind Obedience: Rules, Community and the Individual." In

Meaning Skepticism. Edited by K. Puhl. Berlin: Walter de Gruyter. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

———. 1994a/1999. “Private States and Public Practices: Wittgenstein and Schutz on Intentionality.” *International Philosophical Quarterly* 34: 89–110. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

———. 1994b/1999. “The Significance of Learning in Wittgenstein’s Later Philosophy.” *Canadian Journal of Philosophy* 14: 173–204. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

———. Forthcoming/1999. “The Etiology of the Obvious.” In *Wittgenstein in America*. Edited by T. McCarthy. Oxford: Oxford University Press. Reprinted in *Wittgenstein, Mind, and Meaning: Toward a Social Conception of Mind*. New York: Routledge.

Wittgenstein, Ludwig. 1961. *Notebooks 1914–1916*. Oxford: Blackwell.

———. 1967a. *Philosophical Investigations*. Edited by G. Anscombe and Rush Rhees. Translated by G. Anscombe. Oxford: Basil Blackwell.

———. 1967b. *Zettel*. Edited by G. Anscombe and G. von Wright. Translated by G. Anscombe. Oxford: Basil Blackwell.

———. 1969a. *On Certainty*. Edited by G. Anscombe and H. Nyman. Translated by D. Paul and G. Anscombe. Oxford: Basil Blackwell.

———. 1969b. *Blue and Brown Books*. Oxford: Blackwell.

———. 1975. *Philosophical Remarks*. Edited by Rush Rhees. Translated by Raymond Hargreaves and Roger White. Chicago: University of Chicago Press.

———. 1978. *Philosophical Grammar*. Edited by Rush Rhees. Translated by Anthony Kenny. Berkeley and Los Angeles: University of California Press.

———. 1979a. *Remarks on the Foundations of Mathematics*. Edited by G. Anscombe, G.

von Wright, and R. Rhees. Translated by G. Anscombe. Cambridge: MIT Press.

———. 1979b. *Wittgenstein and the Vienna Circle*. Edited by Brian McGuinness. Translated by Joachim Schulte and Brian McGuinness. Oxford: Basil Blackwell.

———. 1979c. *Lectures, Cambridge, 1932–1935*. Edited by A. Ambrose. Totowa, N.J.: Rowman and Littlefield.

———. 1980. *Culture and Value*. Edited by G. von Wright and H. Nyman. Translated by P. Winch. Oxford: Blackwell.

Wright, Crispin. 1980. *Wittgenstein on the Foundations of Mathematics*. Cambridge: Harvard University Press.

Wurm, S. A. 1972. *Languages of Australia and Tasmania*. The Hague: Mouton.

Wynn, Karen. 1992a. "Addition and Subtraction in Human Infants." *Nature* 358: 749–50.

———. 1992b. "Evidence against the Empirical Accounts of the Origin of Numerical Knowledge." *Mind and Language* 7: 315–32.

———. 1992c. "Issues Concerning a Nativist Theory of Numerical Knowledge." *Mind and Language* 7: 367–81.

Xu, F., and S. Carey. 1996. "Infants' Metaphysics: The Case of Numerical Identity." *Cognitive Psychology* 30: 111–53.

Yallop, Collin. 1982. *Australian Aboriginal Languages*. London: Andre Deutch.