A COGNITIVE APPROACH TO BENACERRAF'S DILEMMA

Thesis format: Monograph

by

Luke V. Jerzykiewicz

Graduate Program in Philosophy

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

THE UNIVERSITY OF WESTERN ONTARIO
School of Graduate and Postdoctoral Studies

CERTIFICATE OF EXAMINATION

Supervisor	Examiners
Dr. Robert J. Stainton	Dr. John L. Bell
Supervisory Committee	Dr. James R. Brown
Dr. John L. Bell	Dr. Daniel Ansari
Dr. Christopher Viger	Dr. Andrew Botterell

The thesis by

Luke V. Jerzykiewicz

entitled:

A Cognitive Approach to Benacerraf's Dilemma

is accepted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Date Oct. 9 09.

Chair of the Thesis Examination Board

Abstract

One of the important challenges in the philosophy of mathematics is to account for the semantics of sentences that express mathematical propositions while simultaneously explaining our access to their contents. This is Benacerraf's Dilemma. In this dissertation, I argue that cognitive science furnishes new tools by means of which we can make progress on this problem. The foundation of the solution, I argue, must be an ontologically realist, albeit non-platonist, conception of mathematical reality. The semantic portion of the problem can be addressed by accepting a Chomskyan conception of natural languages and a matching internalist, mentalist and nativist view of semantics. A helpful perspective on the epistemic aspect of the puzzle can be gained by translating Kurt Gödel's neo-Kantian conception of the nature of mathematics and its objects into modern, cognitive terms.

Keywords: mathematical cognition, philosophy of mathematics, realism, psychologism, functional architecture, conceptualist semantics, cognition, mathematics, Benacerraf, Chomksy, Gödel, Kant.

Acknowledgements

I have accrued a great many personal and professional debts over the years. Let me restrict myself here to acknowledging explicitly only those that bear directly on the writing of this monograph. Above all, let me express my heartfelt thanks to my thesis supervisor, Robert Stainton, whose firm guidance, patient instruction, unfailing support, and kind help over many years have seen me though this project. It's largely by observing Rob work that I have learned what I know of the philosopher's craft. For this I am very grateful. As well, let me extend my thanks to Andy Brook who took a chance on me many years ago and who also taught me how to read Kant slightly askance. My thanks also to the Carleton Institute of Cognitive Science. At the Institute I learned to view cognitive science as an unified endeavour spanning many fields and methodologies, an endeavour with deep historical roots and an open future. Furthermore, I am very grateful to John L. Bell for his insight and wisdom. I regret that I was not able to benefit more from his instruction. His work remains for me a signpost on the horizon toward which I hope slowly to make my way.

I have had the good fortune to study under some remarkable teachers. I would especially like to thank Paul Pietroski, Susan Dwyer, James McGilvray, Stephen Houlgate, Chris Herdman, Marie Odile Junker, Storrs McCall, Philip Buckley, and Angela Cozea. Their ideas and their methods have, in various ways, shaped my approach to research and to writing—sometimes with results they may not have expected or perhaps even approved of.

I'd like to thank Penelope Maddy, Susan Carey, and Ray Jackendoff for brief but extraordinarily enlightening conversations. Each of them has been an inspiration. I have also benefited from the thoughtful criticisms of Georges Rey, Edouard Machery, Mary Kate McGowan, Ileana Paul, Chris Viger, Pete Mandik, Eric Thomson, and Philip Grier.

Let me also acknowledge the many friends and acquaintances off of whom I have bounced ideas over the years. Here I should mention Dan Robins and Sam Scott in particular. Their help and patience with my little obsessions is greatly appreciated.

Of course, I owe a very special debt of gratitude to my parents Janina and Tom Jerzykiewicz. They have been wonderfully supportive of me along the way. It was from my Mom and Dad that I first learned the meaning and the importance of intellectual labour as well as the habits that it calls for. It's impossible to imagine more loving or understanding parents.

Finally, let me express my warm gratitude to my best friend and partner for many years, Catherine Wearing. Her advice and her criticisms have been invaluable. Most of the ideas in this dissertation were first tested in discussion with her. Above all though, I'd like to thank her for sharing with me some very happy years.

I gratefully acknowledge the Ontario Graduate Scholarship Program and the Social Sciences and Humanities Research Council of Canada who helped fund this research.

Contents

List of Tables

List of Figures

0 Introduction

The appearance of an antinomy is for me a symptom of disease.
. . . Whenever this happens, we have to submit our ways of thinking
to a thorough revision, to reject some premisses in which we
believed, or to improve some forms of argument which we used.
— Alfred Tarski

One of the important unsolved problems in the philosophy of mathematics was first articulated in Paul Benacerraf's [1973] paper, *Mathematical Truth*. There, Benacerraf argued that it is difficult (surprisingly difficult, in fact) to offer an adequate *semantic* account of sentences that, to all appearances, express mathematical knowledge, while simultaneously also providing an explanation of our *epistemic* access to their contents. On a standard realist conception, mathematics studies a body of facts—ones not unlike those investigated by the geologist or the chemist, albeit significantly more 'abstract' in nature. Many working mathematicians and philosophers find such a conception attractive. However, it is widely recognized that mathematical realism faces serious difficulties in epistemology: how, short of magic, can our cognitive access to abstract states of affairs be explained? An alternative conception proposes that we view the practice of mathematics as the playing out of a formal game with correct and incorrect moves. Some who consider themselves 'hard-nosed' naturalists opt for this view. To be sure, abandoning the supposition that mathematics studies full-blooded facts does some violence to our commonsense interpretation of the meaning of mathematical expressions. It does however have the advantage of explaining (or, more accurately, explaining away) mathematical 'knowledge,' thereby making the epistemologist's task tractable. This approach encounters its own set of problems. Chief among them is the difficulty of adequately explaining maths' uncanny utility in natural science. Thesis, antithesis, stalemate.

The past two decades have witnessed a steady increase in the interest paid to mathematics (especially arithmetic) by cognitive scientists, including cognitive psychologists and cognitive neuroscientists.[1] As their projects gather momentum researchers are faced with a choice whether or not to interpret the study of mathematics—including algebra, analysis, set theory, and so on—as the investigation of an objectively existing domain. The choice has direct implications for which empirical research programs are taken seriously and which are shelved as *prima facie* implausible. At the moment, the balance is swinging in the antirealists' favour. This, I will argue, is a mistake. My central thesis in what follows is that

[1] See, for example, Butterworth [1999], Carey [2001], Dehaene [1997], Gallistel and Gelman [2000], Gelman and Gallistel [1978], Hauser and Spelke [2004], Spelke and Tsivkin [2001], Wynn [1998], Xu and Spelke [2000]. There have also been a few attempts by sympathetic, naturalist philosophers to apply empirical results and techniques to problems in the philosophy of mathematics. They include Kitcher [1983] and Maddy [1990], and more recently Giaquinto [2001], and Laurence and Margolis [2005].

cognitive science should opt for an unflinchingly REALIST conception of mathematicalia. But in order for this to be possible, ontological realism about mathematics must itself undergo a transformation both with respect to the semantics it endorses and the epistemic picture it relies on. The realist must be able to offer a coherent, naturalist hypothesis regarding the correct solution to Benacerraf's dilemma. What follows is intended as a contribution to that project.

I argue for my thesis in four Chapters. In Chapters 1 and 2, I articulate the kernel of the ontological realist view and show that it is well-motivated. I try to demonstrate however that the standard, platonist elaboration of minimal ontological realism is untenable and so should be abandoned. The positive proposal begins in Chapter 3. There, I retrace some relatively recent Chomskyan arguments in favour of a cognitivist conception of language. A conceptualist account of the semantics of sentences expressing mathematical content based on Ray Jackendoff's recent [2002] work constitutes one half of the positive story. In Chapter 4 I look to Gödel and to Kant for insight concerning the problem of access to mathematical facts. My positive proposal lays out a naturalistic reinterpretation of Kant's [1781, 1783] solution to the epistemic problem presented by mathematics. The reinterpretation is quite substantial insofar as it involves accepting some of Gödel's [1947] criticisms of Kant and marrying the resulting picture with a modern conception of our cognitive architecture. My claim is that this approach, together with the semantics endorsed in Chapter 3, constitutes a viable realist hypothesis regarding the solution to Benacerraf's problem. If that's right then the experimental cognitive scientist working on mathematics ought to choose this brand of naturalist realism over its rivals as the correct background picture on which to predicate her research.[2]

Montréal, Québec
August, 2009

[2]The solution proposed here bears a family resemblance to that suggested in Maddy's [2007] recent book. I embarked on this project several years ago, inspired largely by Maddy [1990], Falkenstein [1995], and Jackendoff [1992b]. Because my views of Kant, Chomsky and cognitive science differ from Maddy's, the realism I end up with is (for better or worse) quite different from hers. Despite these differences, I'd like to think of my project as falling under the umbrella of what she calls 'second philosophy'.

1 Minimal Realism

Do numbers exist? And why?
And must they be so silent?
— Geoffrey Hellman

I'd like to start with the bare basics. In this chapter I introduce a minimal form of ontological realism about mathematics and outline an argument for why we should take it seriously as a prospective standpoint. The bulk of the ensuing discussion is taken up with motivating this argument's premises. This is time well spent. It will allow us to take up several issues of central importance to the philosophy of mathematics, including *inter alia* the purported truth of mathematical propositions, our default commitments concerning natural language semantics, and maths' touted applicability to natural science. Each of these issues will come up again. Moreover, the argument presented in this chapter will itself serve as a point of reference for the remainder of the monograph. The views addressed in subsequent chapters will be attempts to extend it in new directions, to modify some of its assumptions, or to patch its shortcomings.

Definition

In the simplest case, a philosophy P counts as a species of ontological *realism* about e's—where e's can be qualia, possible worlds, physical objects, properties, ethical norms, or anything else—if it meets two jointly sufficient conditions: first, P holds that e's exist (leaving open the question of whether e's are simple or ontologically reducible for now); second, P holds that the existence of e's is not brought about by human agency, whether individual or collective. That there are e's is not due to human labour, rituals, creativity, inventiveness, thoughts, ideas, conceptions, conventions, or ways of speaking. We can think and act as we please; the e's take care of themselves.[3]

This characterization of realism requires two comments. First, the two conditions are jointly sufficient but neither is strictly necessary. There are other ways of defining philosophical realism. Some of these would perhaps substitute a stronger modality claim than *de facto* existence for the first condition. As well, realisms about certain entities—including days of the week, haircuts, propositional attitudes, and perhaps even ethical norms—would want to relax the second condition or abandon it outright. None of this affects my point: whenever the two conditions just named *are* satisfied, a philosophy counts as ontologically realist.[4]

---

[3]To forestall a (rather surprising) misunderstanding let me say at once that I do not take realism to be incompatible in principle with Kant's work.

[4]This definition allows us to include among the realists those philosophers who believe that there are *bona fide* mathematical facts but that there are no mathematical entities. Among them are Hellman [1989] and Putnam [1967, 2006] who reject the existence of mathematical *objects* in favour of a realism about modalities.

My second caveat concerns the omission of an additional requirement that is sometimes mentioned when defining ontological realism. It is occasionally suggested that in order to count as realist, a philosophy P needs to hold that most of what we know about e's is true or approximately true [Resnik 1997].[5] In my view, this addition makes the minimal list of collectively sufficient conditions on realism excessively stringent. Or, at least, it does so in cases where we are able to register the entities in question independently of the descriptions we offer of them. I see no reason to deny, for instance, that those skeptical pre-Copernican cosmologists who held that heavenly bodies existed but that most of what we believed about them was hokum were nonetheless astronomical realists.

If we are prepared to accept the above definition then it is sufficient for a philosophy of *mathematics* to count as realist if it holds that:

($\mathcal{R}$.i) Some mathematical entities exist; and
($\mathcal{R}$.ii) Their existence is independent of human minds, cultures, languages, and conventions.

Let us call a view that accepts both of these clauses, but nothing more, a naked ontological realism about mathematics (NORM, for short). NORM is specific enough to avoid mislabeling as 'realist' any species of nominalism or social-constructivism, since the former violate ($\mathcal{R}$.i) and the latter ($\mathcal{R}$.ii). Nonetheless, NORM leaves unaddressed almost all of the interesting questions that we might expect a philosophy of mathematics to take a view on: the nature of axioms, the role of diagrams in proofs, the bounds of our ontological commitment in 'recreational' mathematics (if any), the reason for maths' apparent indispensability to the natural sciences, and so on. Importantly, NORM is silent on the metaphysical nature of mathematical entities themselves since it says nothing about mathematical entities' essential non-mathematical properties (if any). Some of the philosophical terms are thus left at a pre-theoretical, intuitive level for now. In particular, the mode of existence peculiar to mathematical entities is left as a topic for further study. Finally, because NORM does not give the necessary condtions that must be met for a view to be considered mathematically realist, it leaves the door open to other, independent ways of construing realism. This is deliberate. NORM is no more and no less than one core around which we might construct a realist conception of mathematics. For our purposes, that's what we will need.

Evidence

Whether you prefer to call it parsimony or a healthy skepticism, we should begin from a defeasible presumption *against* the existence of any metaphysical posit. Every new addition

[5]The inspiration here is Dummett. See for instance his [1994].

to our scientific ontology must be warranted.[6] My next task must therefore be to show how our natural skepticism is overcome in the case of mathematical entities. Given what has just been said, there are two distinct questions to answer: Why think that mathematicalia exist? And, assuming that they do, why take them to exist independently of human activity? Let's begin with the first of these questions and return to the second a little later on.

In essence, the argument for NORM's ($\mathcal{R}$.i) runs as follows:

1. Because mathematics is indispensable to the conduct of natural science, some mathematics is *true*.
2. Assuming a standard semantics, the existence of mathematical *truths* entails the existence of mathematical *entities*.

$\therefore$ ($\mathcal{R}$.i) Some mathematical entities exist.

The argument's premises are far from self-evident and a good deal will need to be said to motivate them. We will need to know more about what maths' supposed 'indispensability' to natural science comes to. And it's perhaps not immediately clear what semantics contributes to the issue. Let me try to make this a bit more explicit.

Indispensability and truth

Aristotle's *Metaphysics VI (E)* teaches that "to say about what is that it is not, or about what is not that it is, is false; while to say of what is that it is, or of what is not that it is not, is true" [Aristotle 1941]. Even if somewhat underdeveloped, this characterization of truth is certainly serviceable. Taking it on board we are faced with three broad possibilities regarding the truth value of such assertions as that there is no greatest prime or that first-order Peano arithmetic admits of nonstandard models. First, such statements may correctly describe an existing state of affairs and so be *true*. Alternatively, they may incorrectly describe a state of affairs and so be *false*. Finally, they might not describe a state of affairs at all and, like the configurations of a kaleidoscope, lack content altogether. On this third scenario, apparently true mathematical propositions would be no more than 'acceptable' moves in a complex formal game and so (strictly) neither be true nor false.

Below, I present two arguments. Both start from maths' 'indispensability' to the conduct of natural science, though the precise details in each case are rather different. The first argument discusses maths' role in using accepted scientific theories to *deduce* specific predictions about the behaviour of familiar natural systems and concludes that standard mathematics is not false. The second argument addresses maths' role in helping us extend

[6] As I have already remarked, and as will become increasingly apparent, I take metaphysics to be part and parcel of natural science very much along the lines of Penelope Maddy's [2007] 'second' metaphysics.

our knowledge and *conceive of* new scientific hypotheses, particularly at the murky edges of our understanding. It concludes that standard mathematics does not lack content. Together with what has just been said, these arguments give us a reason to believe that much of what is accepted as standard mathematics is, quite simply, *true*.

Deductive indispensability. At first pass, it may perhaps be natural to think of individual laboratory experiments as capable, under ideal conditions, of directly confirming or disconfirming specific scientific hypotheses. Duhem [1906] famously argues however that, on reflection, this atomistic conception of hypothesis testing is wrong. Even under ideal conditions, predicting the future state of a physical system given its initial conditions is a task which simultaneously involves many physical laws and auxiliary hypotheses. The auxiliary hypotheses can be trivial: that no external forces are operative, that the experimental apparatus is functioning correctly, and so on. Trivial or not though, the logical form of the prediction made by the experimenter is this: if all of the theories (θ) antecedently believed to govern aspects of the system are true *and* all auxiliary presuppositions (Aux) are true *and* the hypothesis being tested (H) is true *then* the predicted result (R) will be observed.

Typically, when a working hypothesis H *is* borne out by experimental evidence we commit ourselves provisionally to its truth as well as to the truth of the network of supporting theses on which it relies. Admittedly, this has the logical form of affirming the consequent (hence the provisional commitment). But while not strictly valid, the move is an inference to the best explanation we are currently able to supply. The inductive case for H is strengthened as it plays a role in an increasing number of successful experiments.[7]

In cases where the predicted result is *not* observed, the experimenter's *modus ponens* is reversed into the corresponding *modus tollens*:

1. If $(\theta_1 \dots \theta_n) \wedge (Aux_1 \dots Aux_n) \wedge H$, then R
2. Not-R

$\therefore$ 3. Not-$\{(\theta_1 \dots \theta_n) \wedge (Aux_1 \dots Aux_n) \wedge H\}$

Thus, if predictions are not borne out by experiments (or whatever tests are appropriate to the domain) it follows that one or more of the corpus of interrelated theses on which one based the prediction is false. But which? The tricky thing—and this is Duhem's point—is that it does not necessarily follow that the problem lies with the hypothesis currently under investigation. A theory-builder has a good deal of latitude when accounting for recalcitrant data. She can reject the working hypothesis. But she can also reject some of the operative

[7] I am bypassing a discussion of anti-inductivism here since addressing it would take us too far off track. So far as I can see, nothing prevents a strict Popperian from reconstructing my point without reference to induction. Cf. Popper [1963].

auxiliary assumptions or cast a critical eye over the background theories. Some of the artistry of theory-building lies in developing a feel for how deep to dig.

Duhem's argument is bolstered and extended by Quine [1953]. Duhem takes holism to be a feature of theory testing in physics and explicitly rejects the idea that similar considerations apply in physiology or in chemistry. In particular, he does not regard the argument as having anything to tell us about logic or mathematics—in part perhaps because he was not aware of the changes to logic ushered in by Frege. Moreover, Duhem takes the antecedent of the conditional in premise one (above) to naturally be limited to a bounded set of 'relevant' considerations. Quine takes a different view on these three issues. He takes evidentiary holism to be a perfectly general feature of scientific investigation and hence to be domain independent. He sees the argument as embracing even the propositions of logic and mathematics. Such propositions are somewhat unusual insofar as they play an auxiliary role in most (if not all) of our scientific theorizing. For this reason, they enjoy an unparalleled degree of inductive confirmation and reside near the centre of our web of belief. Nevertheless, according to Quine, when confronted with sufficiently strange empirical results (as in quantum mechanics, for instance) we are in principle free to contemplate rejecting or altering even our logic and mathematics.[8] Finally, Quine suggests that our entire world-view is put to the test with each experiment, hence that "the unit of empirical significance is the whole of science" [Quine 1953]. This allows him to maintain that no statement, no matter how precious, is in principle immune to revision in light of future experience.[9]

One need not subscribe to all of the tenants of Quinean underdetermination to find the basic insight convincing. We all frequently rely on predictions about future states of physical systems. Experimental scientists and engineers do so daily and explicitly as part of their professional practice. The rest of us do it implicitly every time we set foot in an elevator, cross a bridge, or swerve to avoid a hazard (to name just a few obvious instances). Whether we know it or not, the underlying reasoning takes the form of a *modus ponens*: provided some initial conditions are satisfied, and a given set of laws govern the system, and a number of auxiliary hypotheses hold, then we can expect the system to behave thus and so.

1. If $(Init_1 \ldots Init_n) \wedge (\theta_1 \ldots \theta_n) \wedge (Aux_1 \ldots Aux_n)$, then R
2. $(\theta_1 \ldots \theta_n) \wedge (Aux_1 \ldots Aux_n) \wedge (Init_1 \ldots Init_n)$

$\therefore$ 3. R

The propositions of logic and mathematics play an auxiliary role in many instances of

[8] For a lucid and wide-randing assessment of indispensability arguments, see Putnam [2006]. For a helpful discussion of the contrast between Duhem's and Quine's views more specifically, see Gillies [1998].

[9] Unfortunately, it also allows Quine to maintain that any statement could, in principle, be held true come what may provided that we were willing to countenance sufficiently drastic changes to our conceptual scheme. For a compelling criticism of Quine's position on this matter see Laudan [1990].

deductive reasoning of this kind. (Whether they are semantically empty and therefore eliminable by clever paraphrase is a question to which I will return momentarily.) Now let us suppose for the sake of argument that the propositions of arithmetic, trigonometry, vector algebra, and calculus which enter into these chains of reasoning are, in point of fact, contentful, substantive but *false*. In other words, they predicate something false of an existing state of affairs. If so, we routinely rely on unsound inferences in order to make predictions. Curiously, those unsound chains of reasoning work very well for us. They reliably lead us from our premises to correct conclusions and have done for hundreds of years. Yet if that's so then we have here a bizarre coincidence on a cosmic scale, a coincidence that constitutes a crisis for our most intimate conception of correct reasoning and inference! It would seem then that the hypothesis that mathematical propositions are false leads in a few short steps to preposterous conclusions.

One might object at this point that, in spite of what I have said, standardly accepted mathematical propositions could still be false. We relied on Newton's mechanics for several hundred years before subtle flaws came to light. Reliability—even reliability in the long run—does not guarantee truth *per se*.[10] There's no denying that the the observation is accurate. But I don't think it fatally undermines my point. Consider: mathematics could have, in principle, injected an *additional* source of uncertainty into our scientific deductions. That is, it could have been the case that, in addition to flawed instruments, biased observations, incorrect auxiliary hypotheses, and the myriad other reasons that our predictions sometimes go astray, we might have had to worry about subtle divergences between our mathematics and the world at large. One can readily imagine, for instance, that our attempts to calculate the angular momentum of a body or the refraction angle of a beam of light only yielded correct results in ninety nine percent of cases. In the other one percent, the mathematics simply became "unglued" from the world—indeed, this is what we would expect were our mathematics subtly false. Were the world constituted in this way, investigators would have simply learned to put up with a certain amount of ubiquitous, math-induced 'slack' or 'randomness' in our science.[11] Yet everyone but the radical sceptic should grant that our calculations do not introduce an additional source of noise into our scientific deductions. It's genuinely remarkable that nothing of the sort happens.[12] I want to suggest that the impeccable deductive translucency of mathematical statements is readily explained either by their utter lack of content (as the formalist asserts) or by their truth (as the realist suggests). By contrast, this absence of random misfirings becomes all but

[10] I am grateful to Mary Kate McGowan for raising this objection.

[11] The sort of randomness I am describing would likely be encountered by chimpanzees were they to attempt to develop a version of classical mechanics. Chimpanzees' basic arithmetic abilities can be very good, but (unlike our own) never wholly reliable. Cf. Kawai and Matsuzawa [2000].

[12] A philosopher who denies this last statement must produce *evidence* rather than a bare possibility. Radical skepticism is not an objection.

inexplicable if mathematics is contentful yet false, as the sceptic would have it.

Ampliative indispensability. Discounting the (admittedly farfetched) possibility that mathematical research generally yields contentful yet false results does not yet establish our conclusion. Many philosophers and some mathematicians have been tempted by the view that geometry, algebra, analysis, and indeed all of mathematics comprises a vast formal game—not unlike chess perhaps—though richer and more abstract. On this view, the propositions of mathematics do not add an additional source of error into the deductions of natural scientists since, far from being subtly false, they *lack content* altogether. We proceed by picking axioms and studying their legitimate consequences. Thus, mathematics is not really *about* anything; or, at most, its claims are about the notation itself and nothing beyond it, much the way that the configurations of cat's cradle do not signify anything beyond themselves. Formalism of this kind is tempting as it allows the mathematician to short-circuit apparently idle philosophical deliberations and press on with the business of studying mathematics itself:

> the typical working mathematician is a platonist on weekdays and a formalist on Sundays. That is, when he is doing mathematics he is convinced that he is dealing with an objective reality whose properties he is attempting to determine. But then, when challenged to give a philosophical account of this reality, he finds it easiest to pretend that he does not believe in it after all. [Davis and Hersh 1998, 321]:

Expressing a measure of sympathy for the formalist view, Saunders Mac Lane [1986] writes that "mathematics is not true, but its correct results are certain." He advises us to abandon fruitless discussions about the *truth* of mathematics and the nature of its corresponding 'objects.' It would be more useful, he suggests, to enquire whether some piece of mathematics is *ingenious*, *illuminating*, and whether it promises to open new avenues for future research. Without a doubt, there is a great deal to be said for Mac Lane's sober, deflationary recommendations. Philosophers can surely be accused of the occasional pin-head waltz. Nonetheless, in this case, I think there is convincing evidence that the formalist is excessively modest in her assessment of math's significance; a good deal of mathematics possesses robust semantic content even when our best minds are disposed to suggest otherwise. Indeed, we will shortly see that, in the final reckoning, Mac Lane himself reluctantly concedes that formalism is untenable.

Geometry. The argument for recognizing maths' robust semantic content turns on some very puzzling historical developments that took place in the course of the late eighteenth, nineteenth, and early twentieth centuries.[13] At the start of the period, a tradition stretching back to Aristotle still took mathematics essentially to comprise two unequal fields of

[13]The classic presentation of the view that I defend here is due to Wigner [1960]. The general form of the argument I offer in this section is based on Mark Steiner's 1998, 2005 reworking and elaboration. While I find

study: geometry and arithmetic. The former was concerned with continuous magnitudes, the latter with discrete pluralities. By this time, the development of new branches of mathematics—notably the calculus and Descartes' algebra of species—had started to exert pressure on this two-fold division of the field. But the picture had not yet given way.[14] Since antiquity, geometry had been taken to be the more fundamental science: while all discrete quantities could be represented geometrically, some magnitudes were known to be inexpressible as ratios of whole, discrete quantities. Moreover, thanks to Euclid's systematization, geometric proofs were taken to epitomize rigorous mathematical demonstration. Whatever flaws had been found in the work by the eighteenth century were minor compared to the problems that plagued contemporary proofs in algebra, probability, or analysis.[15] Geometry stood out as a field in which an empirical phenomenon, physical space, had apparently been regimented and made rational.

The eventual paradigm shift had its roots in the ongoing efforts to tidy up outstanding issues in Euclid's work. Imperfections in the *Elements* were already noted in antiquity. It had long been known, for instance, that Euclid occasionally helped himself to certain 'intuitive' results that, strictly, cannot be derived from the axioms.[16] The most serious problem with the *Elements* however concerned Euclid's fifth 'parallel lines' postulate. Here is Playfair's version: given a straight line AB and a point a not on that line, it is possible to draw a unique, coplanar line through a that will never intersect AB no matter how far it is extended in either direction. The postulate seems plausible enough; it is hard to imagine how it might fail to hold. Nonetheless, Euclid had apparently not regarded it as completely self-evident and avoided making use of it until relatively late in his work. Over the centuries, a number of attempts had been made to demonstrate that the fifth postulate was not

Steiner's anti-formalism compelling, I don't think he's correct in seeing his Pythagoreanism as incompatible with naturalism. Steiner [1998], I think, makes a significant false step when he claims that "it goes without saying [that] the naturalist rules out in advance any connection between the human brain and the universe as a whole, except those connections explained away by natural selection" (p.72). I don't think this is obvious at all. In fact, neither ontological nor methodological naturalism need to commit to such an implausibly strong thesis. I take up this issue in subsequent chapters, and especially in Chapter 4.

[14]In Chapter 4, we will discuss the Kantian conception of mathematics which, in its original [1781] form at least, was based on this picture.

[15]In fact, at the time, many mathematicians still relied on 'physical' demonstrations or eschewed rigorous proof as unnecessary pedantry. Even those mathematicians who were convinced of the importance of rigour were hampered by the fact that no commonly agreed formal definition of such basic notions as number or function had yet been found. In part for this reason, negative and complex numbers were highly controversial and not widely accepted.

[16]In Proposition 12, for instance, one assumes that a line passing through the centre of a given circle must intersect it at two points. Strictly speaking, nether the definition of line nor that of circle guarantees this. Indeed, for this reason some seventeenth century mathematicians, including Descartes, had already come to suspect that the classical geometers of antiquity had surreptitiously relied on much more powerful methods than those published in the *Elements* in order to make their discoveries and that they had deliberately hidden this from posterity. See Detlefsen [2005].

independent of the other axioms. These had proven unsuccessful. At most, mathematicians had managed to show that the fifth postulate is equivalent to a number of other interesting but equally tendentious propositions. (Gauss, for instance, showed that the parallel lines postulate is tantamount to the proposition that a triangle of arbitrary size can be drawn.)

Systematic alternatives to Euclidean geometry were first seriously pursued independently by Gauss, by Lobachevsky and by Bolyai. Gauss chose not to publish on the topic, perhaps for fear of ridicule. He may however have indirectly encouraged Johann Bolyai's interest [Eves 1997]. In the end, it was Lobachevsky who first published a systematic treatment of what he called *pangeometry.* In his work, Lobachevsky replaces Euclid's fifth postulate with an alternative: a given straight line AB and a point a not on that line, more than one coplanar line passing through a can be drawn that does not intersect AB. Given this and the remaining standard Euclidean postulates, it's possible to show that at least two lines through a and parallel to AB can be drawn; in addition, there exist many so-called *hyperparallels* that do not intersect AB and pass through the two parallels only at a. At nearly the same time (so in 1854) Riemann proposed another non-Euclidean geometry; one where *any* two straight lines in a plane intersect. Evidently, in such a system, rather than having more than one parallel to AB, we have none.[17]

The initial reception of non-Euclidean geometry was cool—though understandably so.[18] Kline [1972] suggests two reasons for the lack of interest. From a strictly formal point of view, it was not known whether the new systems were internally consistent. No contradictions had yet emerged, but it was unclear whether any would. And there was a second problem also. As I mentioned earlier, the geometry of Euclid was taken by just about everyone to characterize the structure of physical space. Since the new geometries were based (in part) on quite different and incompatible assumptions, they evidently could not be about real spatial relations. They were therefore useless almost by definition. On the whole, it's quite understandable that a potentially inconsistent formal system with no real applications generated little enthusiasm.[19] The situation changed somewhat after the demonstration—by Beltrami in 1868—that the new geometries were indeed formally consistent relative Euclidean geometry. In effect, Beltrami showed that the geometries of Lobachevski and Bolyai could be treated as inscribed on a plane of constant negative curvature, say a tractoid; the geometry of Riemann can be represented on a surface with a constant positive curvature, say a sphere. The existence of these models liberated researchers dedicated to pure mathematics to explore the implications of the new ideas without fear of contradiction. It however did

[17]For a detailed and largely nontechnical treatment of Lobachevskian and Riemannian geometry, see Eves [1997]. The details are not crucial to our concerns and so I will not review them here.

[18]In an 1840 German exposition of his work, for example, Lobachevski regrets the lack of interest in his writings on the part of the mathematical community.

[19]There were some exceptions; Gauss and Reimann both saw that which geometry applied to the space of our experience as an open empirical question. Both considered the topic worth pursuing.

not change the attitudes of those with a more practical bent, as can be seen in this passage by Duhem [1915]:

> Riemann's doctrine is a *rigorous algebra*, for all the theorems which it formulates are very precisely deduced from its basic postulates; so it satisfies the geometric spirit. It is not a *true geometry*, for, in putting forward its postulates, it is not concerned that their corollaries should agree at every point with the judgements, drawn from experience, which constitute our intuitive knowledge of space; it is therefore repugnant to common sense. [Duhem [1915], quoted by Gillies [1998]]

Duhem expresses here a hostility that was prevalent in the nineteenth century and (evidently) persisted into the twentieth. What's important for us though is that pure and applied mathematics had begun to part company. The latter continued to be useful in science, affairs, and industry. The former started to be viewed by its practitioners as a free, abstract expression of the power of the human intellect—a form of creativity distinct in kind from artistic endeavours but allied in spirit insofar as it was not tethered to the empirical reality in which we dwell. In retrospect, we can see that with the discovery of non-Euclidean geometry the earlier paradigm had been abolished: an apparently finished mathematical field was shown to conceal genuinely new insights and this liberated axiomatic study of geometry from concern with real, physical space [Eves 1997].

Let us at this point register a helpful distinction due originally to Mark Steiner [2005]. It concerns the senses in which a mathematical theory can 'apply' to phenomena. Let's say that the application of a mathematical theory R to a domain D is *canonical* just in the case that R was deliberately developed for the purpose of capturing the formal structure of D. Thus, for instance, the canonical application of the positive integers is the enumeration of stable, discrete, physical entities. Plausibly, real numbers apply canonically to spatial magnitudes (but also felicitously apply to continuous magnitudes of other kinds). Vectors apply canonically to quantities with both a magnitude and a direction, such as the velocity of a moving body. Quaternions were introduced so as to facilitate vector algebra in three dimensions. Their canonical application is to the trajectories of bodies moving under the influence of several, non-coplanar forces.[20] And so on. In general, whenever a mathematical theory is developed so as to model a given domain, its success in so doing requires no special explanation. Nor does its success in capturing the structure of closely related, isomorphic domains. On this definition, non-Euclidean geometries are remarkable precisely insofar as they *lack* a canonical empirical application. Nor are they alone in this respect. Let me now briefly outline the development of a second mathematical theory that shares this feature.

[20]Quaternions are of the form $a + bi + cj + dk$ where $\{a, b, c\} \in \mathbb{N}$ while $i^2 = j^2 = k^2 = -1$. For an interesting exposition of Hamilton's struggle to arrive at a consistent algebra for quaternions, see Kline [1980], Chapter 4.

Groups. Roughly the same time-period witnessed a parallel maturation in algebra and the emergence of algebraic theories which were likewise inspired by purely math-internal considerations, rather than by the attempt to model external phenomena. One of the challenges facing mathematicians in the late eighteenth century concerned finding the roots of arbitrary polynomial equations.[21] A way of determining the roots of quadratic equations had been known since antiquity. To find the values of x that satisfy $ax^2 + bx + c = 0$ we can use the formula $x = (-b \pm \sqrt{b^2 - 4ac})/2a$ and locate the results in the complex plane. During the sixteenth century, similarly general, though more elaborate, means were found of solving cubic and quartic polynomial equations. These involved reformulating the equations by the application of a sequence of arithmetic operations and the extraction of roots (square roots, cube roots, and quartic roots). Some degree five and higher polynomials were solvable by extracting roots of unity as well. However, it was shown during the early nineteenth century that no general formula exists for solving quintic or higher degree polynomials. Explaining which of those equations are solvable by radicals and why became a major research problem. The solution and explanation was ultimately arrived at by Galois, who laid the foundations for an important new area of pure mathematical research in the process.

In general, think of a field F as a subset of the complex numbers that is closed under the arithmetic operations—i.e., given any two elements of the field, applying one of the four basic operations yields a member of F. The rational numbers, for instance, form a field. It's possible to associate with an arbitrary polynomial equation $f(x)$ a field K containing its coefficients—the equation's *domain of rationality.* Some polynomials are reducible over their domain of rationality; their roots are contained in K. Thus, $x^2 - 1 = 0$ for example has both coefficients and roots in the field of the rational numbers. Many polynomials however, including quadratics such as $x^2 + 1 = 0$, do not have this property. It's possible to associate with each such function an extension of K, called a *splitting field*, the smallest set containing K as well as the equation's (possibly complex) roots. In our case, since $x^2 + 1$ has two roots, $\pm i$, the splitting field contains the rationals and $\pm i$. Galois went on to analyze the internal structure of splitting fields. He associated with each a *Galois group*; a set of automorphisms for K which permute the polynomial's roots but leave the underlying field unaffected. In the above case, each rotation of the roots in the complex plane about the point of origin by $n\pi$ for an integral value of n has this effect. Galois showed that whether a polynomial is solvable by radicals depends on some the properties of its Galois group. The subsequent study of the structure-preserving substitutions he employed led some forty years later, to the explicit recognition of the modern group concept.

In modern terms, a group is just a set closed under identity, inverse, and a single binary associative operation. To make this concrete, the integers constitute a group under addition; the reals are a group under multiplication; the rotations of an equilateral triangle about its

[21] My discussion here draws especially on [Parshall 2008], Liebeck [2008] and on Kline [1972, 1980].

centre at $2/3\pi$ intervals also constitute a group. Clearly, groups are a very general notion. And they have proven useful in a variety of mathematical domains. Importantly for us, much like non-Euclidean geometries, group theory has no canonical empirical applications. The group concept emerged originally to help study *other mathematical structures*. This fact was not lost either on the mathematicians who pursued pure mathematical research nor on natural scientists and engineers. We see here yet another reason why both camps found it necessary to revise their conception of the nature of mathematical research. At the opening of the nineteenth century it was still reasonable to believe that mathematical propositions were a sort of description abstracted from observations of external, physical states of affairs. By the century's close, that view was widely discredited.[22] As mathematics became ever more abstract, pure mathematicians celebrated the liberation of their research and maths' new autonomy. They were increasingly left to pursue 'free creations of the human mind.'[23] But while 'free,' these creations were not wholly arbitrary or unconstrained. Pure research was guided by certain mores internal to mathematics. These included (of course) formal consistency but also certain aesthetic considerations: as we have already seen, good ideas are those that were ingenious, illuminating, fruitful, mathematically significant, and (above all) those that are beautiful. Here is Hardy [1940] addressing this topic:

> It would be quite difficult now to find an educated man quite insensitive to the aesthetic appeal of mathematics. It may be very hard to define mathematical beauty, but that is just as true of beauty of any kind—we may not know quite what we mean by a beautiful poem, but that does not prevent us from recognizing one when we read it.... The mathematician's patterns, like the painter's or the poet's must be beautiful; the ideas like the colours or the words, must fit together in a harmonious way. Beauty is the first test: there is no permanent place in the world for ugly mathematics. [Hardy 1940]

The new spirit was not always celebrated by more practically minded researchers. Physicists and engineers in particular were not always supportive of maths' newfound freedom. Mathematical domains with canonical application to scientific problems continued to receive attention. More fanciful areas of pure research however were viewed with increasing skepticism. Davis and Hersh [1998] report that

> in 1910 a board of experts including Oswald Veblen and Sir James Jeans, upon reviewing the mathematics curriculum at Princeton, concluded that group theory ought to be thrown out as useless. [p.205]

The attitude was a sensible one. Such qualities as ingeniousness, simplicity and elegance are evidently interpreter-relative. Functions that prove tricky to define in some formal systems

[22] Recall, for instance, Frege's [1953] withering comments about Mill's "gingerbread arithmetic."

[23] The phrase is Dedekind's.

are trivial to state in others—some functions that are hard to write in C^{++} are easy to state in *Prolog*, for instance (though more often the reverse). Aesthetic notions like the beautiful or the sublime are, if anything, even more obviously interpreter-relative. Which structures one appreciates as possessing such qualities depends in some measure on one's cognitive makeup. It would have been unreasonable to expect human aesthetic judgements to carry cross-species—nevermind universal—validity. Certainly, there was no real reason at all to expect concepts developed to suit human aesthetic sensibilities to be of much use in theoretical physics or in any other branch of empirical science.

Subatomic physics. We now know that Veblen and Jeans were wrong. The surprise came when apparently idle mathematical theories—including both Riemannian geometry and group theory—began to play an indispensable role in facilitating discoveries of hitherto unsuspected phenomena in physics. The early 20^{th} century was a time when, Mark Steiner [1998] notes, physics and chemistry encountered a serious and unprecedented problem: researchers interested in atomic and subatomic interactions stopped being able to rely on macroscopic models to make progress. Events at the atomic scale are sufficiently alien that the discrete, slow, medium-sized objects which we experience as our *lebenswelt* fail to provide the theory-builder with fertile ground on which to base new models. A number of the breakthroughs were due to the application of apparently useless mathematics to this field of study.

An early example concerns the discovery of anti-matter. In 1929, Paul Dirac derived equations that combined quantum mechanics and special relativity to describe the motion of electrons in electric and magnetic fields [Barnett et al. 2000]. Surprisingly, the theory predicted the existence of positively charged "holes" or anti-electrons—and these, Dirac initially took to be protons. A year earlier, Hermann Weyl had published his important [1928] study of the application of group theory to quantum mechanics. Not all physicists were persuaded that a mathematization of physical experiments in terms of group theory could be useful. Still, Weyl argued on purely mathematical grounds that Dirac's anti-electrons could not be protons after all; symmetry required that the new particles possess the same mass as electrons but the opposite electric charge. Dirac did indeed change his mind and about a year later the prediction of anti-matter was borne out by Carl Anderson's discovery of the positron in a cloud chamber photograph of cosmic radiation passing through a lead plate. (Both Dirac and Anderson eventually received Nobel Prizes.)[24]

The power of the theory of symmetries to find structure in the subatomic world and (just as importantly) to direct researchers to new findings became indisputable after the discovery of the Ω^- baryon.[25] By the early 1960s, over a hundred subatomic particles had been found.

[24] For the history, see Hanson [1963] as well as Peter Pesic's helpful introduction to Weyl [1934].

[25] My discussion here follows Steiner [2005] as well as Barnett et al. [2000] and Riordan [1987]. I also draw sporadically on Das and Ferbel [2003].

It was hard to accept that they should all be fundamental. In any case, various patterns had been discerned. An apparently basic cut ran between leptons, such as the electron, and hadrons like the proton and neutron. Particles were further classified according to their various properties or quantum numbers. The hadrons thus divided into baryons and mesons according to their angular momentum (or spin). Observed mesons typically had a spin value of zero. By contrast, baryons took half-integer spin values (so $1/2\hbar$, $3/2\hbar$, and so on). Beyond that rough-and-ready classification however things were uncertain. Some subfamilies of particles seemed to have three members, others six, or seven. The values of some properties (electric charge, spin, strangeness) seemed tightly constrained, while others (mass) varied widely. Moreover, not all quantum numbers were necessarily defined for every particle since not all particles are affected by all forces. Leptons are subject to electromagnetic and weak forces as well as to gravity. They are not however subject to the strong nuclear force and hence most strong quantum numbers are undefined for leptons. By contrast, baryons, including the proton and neutron, as well as mesons are subject to all forces including the strong nuclear force. What complicates matters still further is that not all conservation laws apply equally to each family of particles. Hadrons undergoing some weak and electromagnetic interactions fail to conserve some of their quantum numbers (isospin and strangeness in particular). In sum, trying to develop a neat classification that divides the particles into families and predicts the possible results of high-energy interactions was devilishly hard (quite apart from complications introduced by experimental error).[26]

One way of organizing the data was to offer a group theoretic description. There's an interesting link (first elaborated by Emmy Noether) between invariances, or symmetries, and physical conservation laws: whenever a physical quantity is conserved, we can find a symmetry or invariance associated with it. Conversely, whenever there is an underlying symmetry in a physical system, we can define a conserved quantity associated with that invariance.[27] This observation helps constrain the structure of physical theories. For instance, it was known that the angular momentum of the electron could take one of two possible values. Any electron observed will therefore be in one of two possible states: spin up or spin down. (This is, of course, different from the angular momentum of slow, mid-sized objects which varies continuously.) The probability that a given electron will be found in a particular state varies continuously as a function of its rotation. The spin of the electron is $1/2\hbar$, which is to say it takes *two* full revolutions of 360° to bring the particle back to its initial state. We have no macroscopic model of this phenomenon. Nonetheless, mathematical considerations suggest a particular symmetry group, the standard unitary group in two dimensions, or

[26] Indeed, according to one model popular in the early 1960s, the "bootstrap theory," all particles were equally fundamental and could potentially emerge from any interaction [Riordan 1987].

[27] Here are some familiar illustrations of transformations in which a quantity is preserved: spatial translation preserves mass, temporal translation preserves energy, spatial rotation preserves angular momentum, and so on Das and Ferbel [2003].

SU(2), as the appropriate model. And this was helpful insofar as it suggested SU(2) as a classification scheme for the spin of particles. (The nucleons, incidentally, also turned out to have a property, isospin, that could be modeled by SU(2).)

Symmetries can be instantiated by groups of determinate size. As an illustration, think here of perfect Platonic solids. The symmetry that applies to such three dimensional solids allows instantiations of four, six, eight, twelve, and twenty sides. We can be sure however that no perfect solid with, say, fourteen sides will be found. Similarly, taking SU(2) to be the correct representation of particle spin constrained the number of members that families of particles could have. Thus, adopting SU(2) let physicists work backwards; that is, to reason from the mathematics to the physics and to predict the structure particle families in terms of spin and isospin. In general, following through this idea allowed the classification some of the discovered entities. It turned out however that the resulting classification was imperfect (for example, K mesons came in two pairs, not a triplet as the symmetry predicted). The imperfections in the organization scheme suggested that something was still missing [Steiner 2005].

Murray Gell-Mann and Yuval Ne'eman worked (independently) to find a deeper mathematical framework—a higher symmetry—which could help make sense of the facts. Mathematical considerations suggested SU(3), a group of which SU(2) is a subgroup. The approach proved successful. Using SU(3) Gell-Mann was able to arrange known mesons and baryons into a geometric pattern; the eight baryons of spin $1/2$ could be arranged in a cube-like pattern in a three dimensional space based on their mass, electric charge, and strangeness. The seven spin zero mesons could be arranged similarly, though a gap remained. That imperfection in the pattern suggested that an eighth member of the family might exist, leading Gell-Mann to postulate the η° meson. The particle was indeed discovered within a year, lending further support to the group theoretic approach to conducting research. The real breakthrough came however when the symmetry was taken seriously in suggesting more arcane posits. SU(3) does not require families to have eight members. The symmetry is also realized in families of one, three, ten, twenty-seven (and more) members. Gell-Mann's interest was turned to baryons of spin $3/2$. Nine members of this family had been found; the SU(3) symmetry, which was still highly hypothetical, required that there be a tenth. In 1962, Gell-Mann postulated that particle and specified the properties that it ought to possess (a strangeness of -3 and a relatively large mass). The prediction would have been impossible without group theory. Remarkably, the particle, the Ω^{-}, was indeed found in late 1963. Evidently, the approach had proved successful not only in reorganizing known data but (importantly for us) also in predicting the existence of hitherto unobserved and unsuspected entities.[28]

[28]The Ω^{-} is composed of three strange quarks, hence its -3 strangeness value. Historically however the postulation of quarks and of fractional electric charges—while contemporaneous with the use of SU(3) as the

Relativity. Group theory is not the only apparently useless mathematical theory which has helped extend the scope of our scientific understanding. My second example concerns Einstein's work both on special relativity and, later, on general relativity. Early on in his career, certainly before 1920, Einstein took an instrumental attitude toward mathematics. He also repeatedly displayed a distaste for excessive formalism and where possible preferred to work with a sophisticated 'physical intuition' [Corry 1998]. Perhaps partly as a result, the mathematics in Einstein's revolutionary paper on special relativity was "relatively elementary" [Yourgrau 2005]. A number of mathematicians, including David Hilbert and Einstein's erstwhile mathematics teacher, Hermann Minkowski, took an interest in studying Einstein's work. Beginning in 1907, Minkowski devoted his energies to re-articulating Einstein's ideas in more mathematically elegant terms. The result was the embedding of special relativity in a four-dimensional, non-Euclidean, space-time manifold. This formulation, whose clarity surpassed Einstein's, was immediately adopted by the first textbook treatments of special relativity. And so, Minkowski space-time became a standard feature of the theory. What should strike us as odd however is that Riemannian geometry, a mathematical theory with no canonical application, should have been useful for this purpose at all. Minkowski did not need to develop a new formal system on the basis of Einstein's work in physics; he did the reverse: he employed ideas that had *already been found* by the pure mathematician in order to clarify what the physicist was laboriously trying to express. Commenting on this, Minkowski says the following:

> To some extent, the physicist needs to invent... concepts from scratch and must labouriously carve a path through a primeval forest of obscurity; at the same time, the mathematician travels nearby on an excellently designed road.... It will become apparent, to the glory of mathematics and to the boundless astonishment of the rest of mankind, that the mathematicians have created purely within their imagination a grand domain that should have arrived at a real and most perfect existence, and this without any such intention on their part. [Minkowski (1915) quoted by Pyenson 1977.]

Einstein himself was not impressed with the reformulation of special relativity in terms of four dimensional, Riemannian geometry; indeed, he considered it 'superfluous erudition.' One reason for this has to do with discrepancies between Minkowski's formulation and observed physical phenomena; Minkowski had, in effect, made some errors [Pyenson 1977]. However, Einstein's estimation of the role of mathematical formalism in research soon began to change. Once he turned his attention to developing the general theory of relativity, he found that some of the formal elements of Minkowski's work (for instance, the invariant line of Minkowski's space-time) came to play an important role in extending his research.[29]

---

classificatory framework for mesons and baryons—was logically independent of it. I am omitting a discussion of quarks here. See however Steiner [2005] and Riordan [1987].

[29] See Corry [1998] for a discussion.

Gradually, Einstein came to appreciate that his own creative efforts in physics depended importantly on the antecedent existence of an appropriate set of mathematical ideas. By the early thirties, he had come to adopt his teacher's perspective:

> Our experience hitherto justifies us in believing that nature is the realization of the simplest conceivable mathematical ideas. I am convinced that we can discover by means of pure mathematical constructions the concepts and laws connecting them with each other, which furnish the key to understanding natural phenomena. Experience remains, of course, the sole criterion of the physical utility of a mathematical construction. But the creative principle resides in mathematics. [Einstein (1933) quoted by Corry 1998]

The development of general relativity thus indirectly depended on Minkowski's work on space-time. And that, in turn, depended on Riemann's work on non-Euclidean geometries. If we construe mathematical systems with no canonical empirical application as literally contentless formal structures, this sequence of events becomes all but impossible to explain. The fact that we are able to form ideas which enable us to limn the structure of reality *at all* is a nontrivial gift. That we can do so while apparently pursuing idle speculation and in advance of empirical investigation is truly remarkable. Certainly, Einstein was struck by the peculiarity of our situation: "the most incomprehensible thing about the world," he remarks, "is that it's comprehensible."[30]

Discussion. It will not do to claim that using non-Euclidean geometry or group theory in physics is on a par with constructing a tool for one purpose and subsequently finding that it can be put to other uses.[31] This sort of deflationist analogy works well to explain why (say) the real numbers can be used to quantify physical lengths as well as temporal durations; both are continuous magnitudes. And the canonical application of the real numbers are precisely such magnitudes. The deflationist analysis works less well in the two cases we have just considered. Neither non-Euclidean geometry nor group theory has canonical empirical applications. Thus, rather than likening them to a tool, it would be more accurate to liken them to a board-game or a Kandinsky painting that, years after its creation, has been found to embody a useful roadmap to a newly discovered continent. Expecting our games or our artistic creations to reflect the structure of far-away lands is like expecting magic to work [Steiner 1998]. Or, rather, it *would be* like relying on magic were we to assume that such formal systems carry no more content than chess configurations or kaleidoscopic images. For just this reason, Saunders Mac Lane is moved to reject this supposition:

> How can one account for the unreasonable effectiveness of Mathematics in providing models for science and knowledge?... To fully account for this applicability, the phenomena must in some sense be ready to fit the formulas. This becomes a question about

[30]Quoted in Yourgrau [2005].

[31]This is suggested by Kline [1980].

> the character of reality. All the experience of Mathematics and of the physical sciences shows that many aspects of reality can be measured, organized, and then understood by way of theories and concepts which have a large formal content. We have to agree that the form is chosen to reflect the facts. It must also be the case that the facts accept the form. We have not explained why this is so... For a strict formalist, for whom Mathematics is just manipulation of symbols, this question cannot be answered. [Mac Lane 1986, p. 445]

There is an alternative to formalism: we can instead accept the notion that a great deal of pure mathematics is contentful *even when mathematicians themselves don't really think that it is.*[32] Such a supposition allows us to hypothesize that a well-hidden structure-preserving mapping exists between certain basic mathematical facts and physical reality. Once we grant that this is the case, the supernatural aura surrounding unexpected applications begins to dissipate. Admittedly, we are left to explain the concrete nature of the apparent preexisting harmony between the physical world and our mathematical discoveries. But this is to trade in an apparent miracle for a legitimate research problem—surely a trade worth making. In any case, we have little choice. Unless we are prepared to countenance a mystery, we must not conclude that mathematics is merely a meaningless, formal game; the notation gives back more than we apparently put in.

Semantics and entities

So far, I have argued that the propositions of mathematics are not purely formal (or content-free) and that at least some of them are not contentful yet subtly false. Given the alternatives, these arguments show that some mathematical propositions are, in fact, *true.* Notice however that it takes an additional argument to move from the truth of mathematical statements to the existence of mathematical entities. The argument I will rely on here takes as its point of departure one of the leading available accounts of natural language semantics.

Formal semantics. Semantics, understood as a branch of linguistics, aims to offer a theory of meaning for natural, human languages. Having a working theory of this sort would help explain how we manage to interpret linguistic expressions, hence how linguistic communication is possible. In order to build a semantic theory it's useful to idealize somewhat by abstracting from extraneous details: the *core meaning* of a sentence, phrase or word is that aspect of its semantic value that remains invariant under context change. Focusing on core meaning allows us to exclude from consideration peculiarities of semantic valence contributed on a given occasion by the identity of the speaker, the time and place of utterance, the gathered company, and other purely circumstantial considerations. It's not wholly

[32] Concretely how much "a great deal" turns out to be is a topic that I will address in a moment.

self-evident how to go about developing a responsible, scientific account of core meaning. Much is owed in this area to the work of Richard Montague [1974b], who himself built on insights due to Frege [1953] and Tarski [2001]. A recurring theme in Montague's work is that, in order to arrive at an appropriate theory of meaning for natural languages, we should treat these as essentially on par with artificial structures, such as the predicate calculus:

> I reject the contention that an important theoretical difference exists between formal and natural languages. [Montague 1974b]

Making this assumption has the benefit of enabling us to employ the machinery of model theory to construct our semantic account.[33] At its core, furnishing a semantics for a *formal* language involves specifying a systematic mapping from its well-formed formulae to an extralinguistic domain comprising at least two elements (say, 1 and 0, or $\top$ and $\bot$). In effect, working out the interpretation of a well-formed formula involves deriving its truth value given a specific assignment of variables to domain elements. Something similar can be done for the sentences of a natural language. We can't identify the meanings of natural language sentences with their truth values *simpliciter*, since that would make two of every three sentences synonymous. Rather, we get our theory off the ground by supposing that the core meanings of natural language sentences are importantly bound up with their truth conditions [Davidson 1967, Tarski 2001]. The idea is quite intuitive. Consider:

(1) "Śnieg jest biały." is true if and only if snow is white.

In order to grasp the core meaning of the sentence named in the antecedent, it suffices to learn under what conditions that sentence comes out true. The appropriate truth conditions are given in the consequent of the biconditional. Generalizing the point, a good first pass at a semantic theory for one portion of a natural language—consisting of the set of its well-formed, disambiguated, present-tense, declarative sentences—consists of an exhaustive list of the mappings from the names of these sentences to the appropriate truth conditions.

There needs to be more to an adequate semantic theory however. We want our theories to be elegant and economical. A theory consisting of sentence-fact pairings is neither. The number of distinct, grammatical, declarative sentences of a natural language is, in principle, unbounded.[34] To fully interpret a natural language an exhaustive list of pairings would therefore need to contain an infinite number of distinct clauses. Not only would it not be elegant; it could never, in principle, be complete. Worse still, human beings do, as a matter of fact, interpret their natural languages on the fly. It's not plausible that we do so by searching an infinite or even a very long list (if only because to do so would be impossible in real time provided we assume a small but fixed processing time for each entry). There is

[33] For a helpful exposition of the semantics of predicate calculus, see Bell et al. [2001] or DeLong [2004].

[34] The grammaticality of arbitrarily many centre-embedded clauses shows this for English.

another problem too. Not all of the declarative sentences whose truth *conditions* coincide are, in fact, synonymous. Think here of the truths of logic or mathematics. Pairing sentences to truth conditions is a good start for a semantic theory, but it's not the whole story. If we want our theory of meaning to be an explanation of how we understand language, we can't stop yet.

To make progress we once again borrow from the semantics of formal languages. The interpretation of complex formal formulae is carried out recursively. One begins by offering an interpretation of atomic components and moving to higher degree cases via recursion over the quantifiers and operators. The semantic value of a complex formula thus turns out to be a monotonic function of the semantic values of its parts (and their arrangement). Essentially the same idea can be used to deal with natural languages. To get the analysis of constituents going we introduce the notion of a semantic type. We say that sentences are of semantic type $\langle t \rangle$, for truth-conditions. Next, we map the various phrases to their appropriate referents. In contrast to the formulae of the predicate calculus, the sentences of a natural language are used to convey information about our surroundings—what Quine calls 'the passing show.' The extra-linguistic domain D to be employed in interpreting natural languages is therefore not merely the true and the false but rather the world itself. Names are mapped directly to the relevant individuals in D. This won't work for common nouns, such as 'cat' or 'dog'. Those we take to pick out sets, such that the extension of a common noun is the set of (actual or possible) entities in D named by the term. For the sake of clarity, we say that nouns are of semantic type $\langle e \rangle$, for entity. A subset of the derived types allows us to classify other linguistic items including logical connectives, verbs, adverbs, and adformulas. Consider first the predicates of natural languages, including verbs and adjectives. They have unsaturated semantic values which require arguments to constitute fully-formed thoughts. It's natural (after Frege) to take them to denote functions. In the simplest case, the case of intransitive verbs, these are functions from entities to truth conditions. That's because such verbs take entities of type $\langle e \rangle$ as arguments, and return semantic types $\langle t \rangle$ as values. Intransitive verbs are therefore naturally classified as of semantic type $\langle e, t \rangle$. The case of transitive verbs is only slightly more involved: they take both a direct and an indirect object. Hence, they take two $\langle e \rangle$'s to a $\langle t \rangle$: so are either of type $\langle \langle e, e \rangle, t \rangle$ or $\langle e, \langle e, t \rangle \rangle$. Indeed, an unbounded range of semantic types can be specified recursively: we stipulate that if σ and τ are types then $\langle \sigma, \tau \rangle$ is also a semantic type. By defining an unbounded variety of semantic types in this way, FS offers us a maximally flexible classification of possible denotations relative to a domain D. Furthermore, there is no need to expand our domain in order to accommodate functions, since functions can be identified with sets of ordered pairs. Thus, a function that maps to $\top$ just in the case that it takes elements of a specific subset of D as inputs can itself be construed as a subset of D. Thus, the entire interpretation M requires nothing more than a domain and an interpretation function.[35]

FS is an elegant theory.[36] It is capable of building up plausible interpretations of an unbounded number of sentences from a finite base. It analyzes the core meaning of a sentence both with reference to its truth conditions and the meanings of its parts, thereby explaining why not all sentences with identical truth-conditions are synonymous. Finally, formal semantics hooks language up with the world, thereby helping make sense of how human beings use it to convey content. Plausibly then, it can serve as the beginnings of a theory of meaning for natural languages.

More could be said about formal semantics; and indeed, I will return to the topic in later chapters. My aim for the moment is to say enough to motivate the theory and to explain how it approaches semantic analysis. Let's therefore put FS to use in the context of mathematical statements. As our example, consider this bit of mathematical trivia: 17 happens to be a prime number. The core meaning of the sentence that expresses this fact is given by its truth conditions in the obvious way:

(2) "Seventeen is a prime number." iff seventeen is a prime number.

The full interpretation of the sentence named in the antecedent requires us to explain how its truth conditions are a function of the meaning of the parts. The first step is to pick out the sentence's parts unambiguously. A very natural way to do that is to use its syntactic description.[37] Here it is:

(3) $[_{\text{S}}\ [_{\text{NP}}$ **Seventeen**$]\ [_{\text{VP}}$ **is** $[_{\text{DP}}$ **a** $[_{\text{NP}}\ [_{\text{AP}}$ **prime**$]\ [_{\text{NP}}$ **number.**$]\ ]\ ]\ ]\ ]$

And here is how those parts are analyzed:

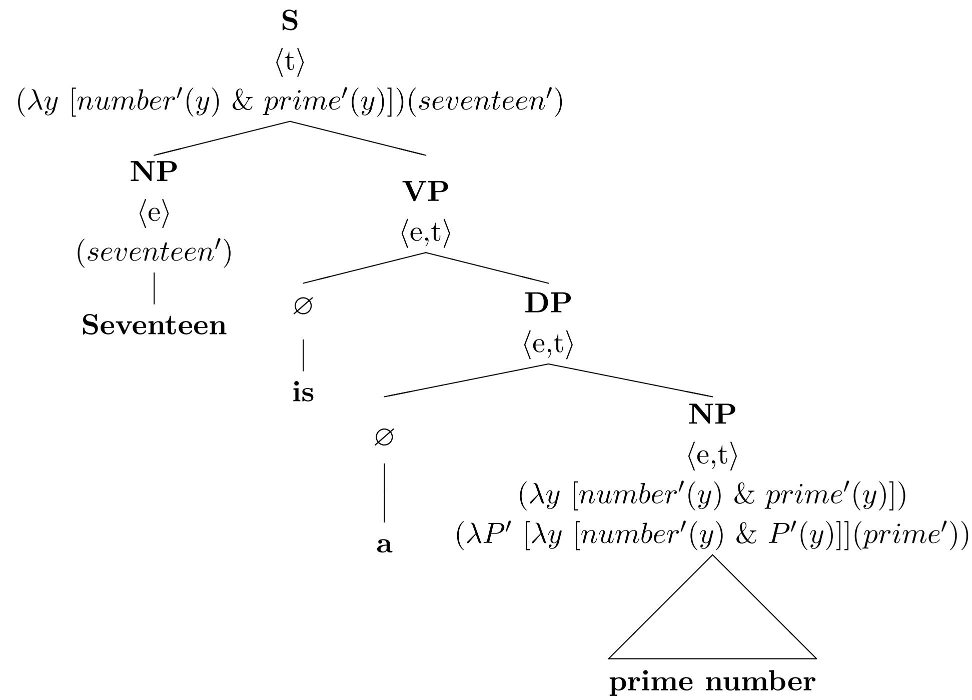
---

[35] Specifying precisely *how* large the domain presupposed by FS needs to be presents certain problems. Certainly, if proper classes exist, they can be referred to. And so on up the iterative hierarchy. It would be tempting to identify D with the iterative hierarchy, taking actual and possible physical items to comprise the bottom layer. The wrinkle is that the iterative hierarchy itself figures in D as a possible object of reference. As we shall see in Chapter 3, this is the least of our worries.

[36] For a useful exposition of Montague [1974b], see Thomason [1974]. For a classic textbook treatment, see Heim and Kratzer [1998].

[37] Interestingly, Montague himself had little time for "developments emanating from the Massachusetts Institute of Technology" and did not employ what are now accepted syntactic categories. We can however follow [Larson and Segal 1995] and Heim and Kratzer [1998] instead.

(4)



Notice that the VP constitutes an unsaturated function from entities to truth values. Both the copula and the indefinite article are semantically inert, contributing nothing to the overall interpretation. The two NPs require a few words of explanation.[38] The final NP is of semantic type $\langle$e,t$\rangle$, and designates a function from entities to truth values. (I am treating ‘prime’ as an intersective adjective since not everything that is prime in the relevant sense is *ipso facto* a number. Primeness has been generalized to other branches of mathematics. In fact, this helps simplify the analysis considerably.)[39] For the sentence to be of semantic type $\langle$t$\rangle$, the initial NP must be of type $\langle$e$\rangle$ and to designate an entity. The whole is therefore true if and only if there in fact is an entity, picked out by ‘seventeen’ which has the property of being a prime number. Ergo, at least one number exists.

And so we arrive at our conclusion. Suppose we accept standard semantics. Suppose also we accept that all sentences displaying a given syntactic structure be treated uniformly by our semantic theory. If so, then the existence of mathematical truths entails the existence of mathematical entities just as the existence of nontrivial facts about London entails the existence of that city. We must either accept NORM’s ($\mathcal{R}$.i) or give up on formal semantics

[38]This reading of the indefinite article in the predicative position, I have been informed, is not uncontroversial, but it is quite standard. My reading is based on Heim and Kratzer [1998]. I am grateful to Rob Stainton, Marie Odile Junker, and Ileana Paul for helpful discussions, suggestions, and criticisms of my linguistic analysis at various stages of this project. The remaining errors are mine alone.

[39]For a helpful discussion of nonintersective adjectives, see Cann [1993], Chapters 6 and 10.

and supply some alternative theory of meaning.[40]

Objectivity

Recall that NORM comprises two theses. We now have a reason to accept the first: that some mathematical entities exist. Before ending the chapter, I want to motivate the second thesis: that *at least some* mathematical entities—whatever they may ultimately prove to be—exist independently of human conventions, opinions, thoughts and activities. They are, in a word, *real*. Here, I should begin by deflating expectations slightly. Because the scope and the metaphysical nature of the mathematical domain remains unspecified, the responsible course for a would-be ontological realist is to start with maximally conservative commitments. And so this is what I will do.

Constructivism. Brouwer's [1975] intuitionism (or constructivism) is perhaps the most influential modern *anti*-realism *within* mathematics.[41] Davis and Hersh [1998] estimate that, currently, one in twenty working mathematicians is a constructivist. There are also related projects in logic, semantics, and metaphysics. It may seem at first glance that in order to defend NORM, one would need to attack Brouwer. I don't think this is the case however. In fact, the very weak sort of ontological realism that interests me here can remain neutral with regard to the dispute between constructivist and nonconstructivist mathematicians. To show that this is the case, I want to clearly distinguish core constructivist commitments regarding the ontology of mathematics from associated peripheral, phenomenological doctrines. My approach will be to concede as much of the mathematical core of constructivism as possible while showing that realism (in my sense) still holds—and hence the bundle of cognitive and epistemic problems that interests me endures. This strategy will force us to restrict the scope of the ensuing discussion to what many mathematicians would regard as an uncomfortably small corner of the mathematical universe. Nonetheless, the compromise has the advantage of temporarily forestalling unhelpful disputes.[42]

It's sometimes said that constructivism eschews independently existing mathematical objects altogether or that it construes mathematical reality as essentially consciousness- or mind-dependent. There is certainly textual support in Brouwer's [1975] work for such attributions.[43] Here however, I will follow George and Velleman [2002] in suggesting that

[40]Notice, by the way, that function-composition seems indispensable to articulating a semantic theory. This reinforces the point I made earlier about the indispensability of mathematicalia in natural science.

[41]I will not have much to say here about more *philosophical* anti-realisms regarding the ontology of mathematics, such as fictionalism. See Stanley [2001] for a discussion and criticism of fictionalism.

[42]The compromise still allows us to work with significantly more of the mathematical domain than is currently under discussion in much more ambitious cognitive works, such as Butterworth [1999].

[43]See for instance Brouwer's enigmatic [1948] essay on *Consciousness, Philosophy, and Mathematics*, in his Collected Works [1975], pp. 480-494.

those speculative metaphysical doctrines are not ultimately what lies at the heart of modern constructivism. It seems to me that it's *au fond* mathematics and not metaphysics that the constructivist cares about. Thus Heyting [1956]: "We have no objection against a mathematician privately admitting any metaphysical theory he likes, but. . . we study mathematics as something simpler, more immediate than metaphysics."

At its core, constructivism is animated by a skepticism concerning the *determinateness* and *completedness* of mathematical reality. A good illustration of this comes from the constructivists' interpretation of Cantor's diagonal proof. Recall that two sets are equinumerous just in the case that a one-to-one, onto mapping between them can be established. The integers and the rationals are both of the same cardinality as the natural numbers since a bijection between them and the naturals exists. Cantor shows that, by contrast, any scheme for establishing such a correspondence between reals and natural numbers is bound to fail. It's always possible to construct a new real number not accounted for by any proposed enumeration. From this, a classical mathematician concludes that the infinite set of natural numbers is smaller than the infinite set of the reals. The former comprise a countably infinite set of cardinality $\aleph_0$ while the latter comprise a nondenumerable set of cardinality $2^{\aleph_0}$. The intuitionist is more circumspect. She accepts Cantor's proof for the non-denumerability of the reals but takes the peculiar classical gloss on this result to be driven by a dubious ideology. In its stead, she offers a substantially different interpretation. The crucial point in contention is the notion of an actual, completed infinity. Nowhere in the proof does Cantor demonstrate that such collections truly exist (nor does he aim to). Indeed, it's not altogether clear how to make sense of the literal existence of completed infinite collections without resorting to fairly fancy metaphysical footwork.[44] What *is* clear is that, given any collection of natural numbers, no matter how large, it's always possible to generate a new natural number not yet on the list (for example, by taking the successor of the largest member of the collection). In view of this, the constructivist proposes that we view the set of natural numbers as (not literally *infinite* but rather) as *infinitely extendable* [Dummett 1994]. Viewed from a constructivist standpoint, Cantor's proof shows that given any rule for extending the collection of real numbers into a more inclusive collection, it's possible to find a real number not covered by that rule. The reals are infinitely extendable, but in a manner that is different in kind from the way that the naturals or the integers are. The example illustrates nicely the considerable overlap between the classical and constructivist perspectives: Both accept the existence of mathematical entities—including numbers, collections, and functions.[45] They merely interpret these differently. In the final analysis

[44]For sample steps from the metaphysician's repertory, cf. Balaguer [1998]. More on this in the subsequent chapter.

[45]Though not always the same collections and functions.

then, the fundamental point of contention concerns not phenomenological issues (the mind-dependence of mathematics, the role of consciousness, or anything of the kind) but rather whether such entities can, in some way, be grounded in completed infinities.

The constructivists' rejection of completed infinities and of a finished, determinate mathematical realm has far-reaching implications. Let me point out just two. My first illustration crops up in logic. Classical logicians assume the existence of freestanding infinite collections. As a result, they rely on arguments that do not take care to distinguish between reasoning in finite and in infinite domains. It's trivial, for example, to divide a *finite* collection into two subsets, one of which contains all and only members that display a given property, and another whose members do not. It's natural to extrapolate from this and to suppose that, in the general case, $p \vee \neg p$ holds true.[46] That is, any proposition whatsoever is either true or false. Were this the case, it would be exceedingly useful since it would allow us to prove the truth of any proposition by demonstrating that its contrary does not hold:

$$\frac{\neg\neg p}{\therefore p}$$

The constructivist objects both to the unbounded use of the *excluded middle* and to proofs by *reductio ad absurdum* in infinitely extendable domains. To see why, consider whether winning the lottery is part of your life-story. Since your life is essentially incomplete, unless you have in fact already won the lottery, this question cannot be settled. It literally doesn't have an answer just yet. Now, admittedly, one could stipulate that everyone's future is fixed and so (in some sense) 'exists' already. But it's far from clear why we ought to find such stipulation scientifically responsible or philosophically compelling. The constructivist is moved by parallel considerations to suspend judgement on such matters as whether the decimal expansion of π contains a sequence of seven 7s. Not only do we not know the answer, she argues, there may well not yet *be* an answer. In general then, it's actually not the case that all propositions just are either true or false. For this reason, neither the principle of excluded middle nor demonstrations by *reductio ad absurdum* enjoy blanket legitimacy.[47]

The second important implication of adopting a constructive conception of mathematics that I'd like to mention arises in set theory. Here, the constructivist is forced to give up

[46] Or, more accurately, that $(x)(P)(Px \vee \neg Px)$ holds.

[47] There's an important subtlety here that I'm glossing over. The reference to unknown future contingents may make it appear that some n-valued logic might provide a suitable vehicle for capturing the idea that the constructivist is articulating. If that were so, constructivist logics would turn out to be a special case of Łukasiewicz's three valued logic. This is not so however. Kurt Gödel showed that no n-valued logic which preserves the classical truth tables can serve as an appropriate semantics for constructivist logic (assuming that the value of n is smaller than the number of entities in the logic's domain). That's because some pair of atomic sentences A and B must take the same intermediate truth value and therefore $A \rightarrow B$ will map to $\top$. But since the constructivist interprets $A \rightarrow B$ to mean that given a proof of A we can derive a proof of B, she will sometimes be forced to deny the truth of the implication. See [Bell et al. 2001, 195] for a discussion.

the unrestricted application of axiom of choice, another one of the classical mathematicians' important tools.[48] Consider: A relation is typically identified with a set of ordered pairs. A function is a special case of a relation H such that for every x in the domain of H, there is a unique y such that xHy. A classical mathematician accepts that given any relation R, there is a function F over the same domain, such that $F \subseteq R$. In the case of finite sets, carving out a function F from a relation R then is just a matter of choosing one (perhaps among several possible) mappings to R's range for each member of the domain. This can be done unproblematically even for infinitely extendable domains consisting of sets of discernable elements. To modify an example from Russell, if our domain consists of denumerably many triples comprising a spoon, knife and fork, we can stipulate by description that our function return (say) the spoon from each set. If however R's domain consists of $\aleph_0$ triples of indistinguishable spoons, there is no way of defining an appropriate function by describing what it should pick out. In this case, the classical mathematician stipulates that there exists a function of the required type. This is legitimate on the assumption that all functions enjoy a freestanding existence regardless of whether we can specify them by description. But this is, of course, just what the constructivist doubts [Posy 2005].

There is no question that adopting a constructive conception of mathematics changes one's understanding of the field. From a classicist standpoint, constructivism constitutes an unacceptable restriction on the freedom of mathematical research [Shapiro 1997]. Without the axiom of choice, excluded middle and proofs by *reductio*, a good deal of what is viewed by many as standard mathematics becomes inaccessible.[49] This is not a negligible sacrifice. Yet from a constructivist standpoint, the classical mathematician runs a serious risk of working with illusory posits—ones that have either not been proven consistent, or worse yet, ones that—in spite of our creative imaginings—do not really exist.

In what follows, I propose to remain strictly *agnostic* regarding the constructivist/classicist controversy. My reasoning is as follows: in general, a scientific realist is provisionally committed to the existence of whatever posits are generally recognized in a given field of study [Melnyk 2003]. One rough and ready way of determining what counts is a survey of standardly accepted undergraduate textbooks. In a domain where serious scientific controversy still reigns, it's correct for a non-participant to suspend judgement. I take the classicist/constructivsit controversy to be an unsettled issue between rival paradigms. It would be irresponsible and uncharitable to pretend to resolve the dispute by fiat. Anyway, I'm not remotely qualified to do that. For this reason, in what follows, I will take NORM to be provisionally ontologically committed to those (and only those) mathematical structures which *both* the classical mathematician and the constructivist recognize. The constructivist

[48] Actually, Gödel showed that the axiom of choice holds in the constructible universe.

[49] I leave a discussion of the constructivsts' rejection of impredicative definitions and discontinuous functions for another time. These are interesting and important issues but discussing them here at length would not further my argument.

rejects the existence of a determinate, finished mathematical universe. On the other hand, the classicist rejects Brouwer's free choice sequences. So be it. I am prepared to limit my ontological commitment to entities that feature in mathematical assertions that hold universally, in every mathematical framework, whether it be constructivist or classical.[50] This has the effect of making NORM a much narrower hypothesis than other proposed ontological realisms about mathematics. I don't see how this can be avoided; nor does this seem like a bad thing given the scientific goals of this dissertation.

Intersubjective conventions. There is another kind of anti-realist view that I'd like to briefly address before concluding the chapter. We have already encountered Hardy's [1940] claim that a certain aesthetic sensibility plays a significant role in guiding the course of mathematical research. Hardy himself was a realist about mathematics and its entities. But one could push his ideas concerning the importance of the subjective element in mathematics quite a bit further than he was prepared to do. Emphasizing the active, constructive role of the mathematician and her intellectual community in the process of mathematical development, one could arrive at a characterization of mathematics—and particularly of pure mathematics with no canonical empirical applications—as an essentially arbitrary, negotiated cultural product. Full blooded social conventionalism would have it that the justification for mathematical claims rests ultimately on intersubjective consensus among maths' producers, the mathematicians (and perhaps also on the entrenched practise of respecting the law of non-contradiction and other such traditional notions). On such an account, it's ultimately aesthetic, psychological, sociological and (ultimately) economic factors that determine what maths' producers find acceptable and compelling. Hence it's to those factors that we need to look so as to discover the nature of mathematical truth.

The radical conventionalism I have just sketched is, of course, a caricature. Not even those who talk about mathematical 'acculturation', 'cultural production', and its eventual fetishization argue explicitly for the theory when it's put in such stark terms.[51] Watered down versions can however be found. In fact, even those theorists who are critical of conventionalism sometimes find themselves inadvertently moving in that direction in spite of themselves. Consider, for instance, Ray Jackendoff's important and highly original cognitive

[50] In Chapter 8 of his [1988] book, John L. Bell asserts that such invariant mathematical laws "are the theorems of the free naturalized local set theory." I expect it will take me some time to understand the concrete content of this claim. The suggestion is worth pursuing.

[51] See, for instance, A. J. Bishop [2002]. According to marxist theorists—for instance Walter Benjamin—a cultural product is fetishized when the labourer forgets its human origins and mistakes it for a sui-generis entity. The notion of *fetishization* is often a centre-piece of marxist critiques of commodification and religion. See Buck-Morss [1991].

theory of concepts.[52] Jackendoff [2003] recognizes and discusses the importance of mathematical objectivity. Nonetheless, his theoretical apparatus leaves him with few resources to stave off a slide into conventionalism. It's useful and instructive to see why this occurs and also why the pull of anti-realism should be resisted.

Concepts, in the sense studied by cognitive psychologists, cognitive neuroscientists and (many) linguists, are a subclass of structured, information-bearing states of an individual's mind/brain.[53] Human brains share many aspects of their functional organization with the brains of nonhuman animals. Our sensory modalities and our basic affective systems are, for instance, a shared biological endowment. Much of our cognitive flexibility and hence our effectiveness as a species has been attributed to our ability to integrate information across a variety of cognitive domains [Spelke 2002]. Jackendoff [1992a] argues that concepts play a crucial integrative role in human mental activity. He construes them as informational interfaces between lower, task-specific or input-specific modules: linguistic, visual, auditory, haptic, and olfactory information, as well as information about the position and state of one's body and one's affective state is centrally encoded by conceptual structures. The integrative work of the conceptual faculty allows us to bind various features of diverse stimuli into a single, coherent representation: we can listen to a piano concerto, individuate *its* movements, watch *it* being performed, remember the last time we had heard *it*, and so on. Apart from integrating input, the conceptual faculty also plays a role in facilitating complex, deliberate behaviour. Finally, what we experience as imagining, remembering, and thinking is—on this account—the result of the largely unconscious recombination of conceptual representations.[54]

I will return to a more sustained discussion of Jackendoff's theory later on. What interests me at the moment is the account of mathematical reasoning implicit in his work. According to the view, conceptual representations of concrete objects display both perceptual and inferential links. The perceptual links allow one to identify the relevant objects and to recall their physical and spatial properties. Inferential links, by contrast, allow one to reason about those objects. The theory makes slightly different provisions for abstract 'objects' such as mortgages, interest rates, and lies. Since such entities have no defining perceptual properties *per se*, they do not link to the sensory modalities; their role in the cognitive economy is exhausted by their inferential links. We learn how to think about

[52] In this section, I limit myself to discussing just one aspect of Jackendoff's work: his theory of concepts. But, in fact, Jackendoff [1983, 1987, 2002] does vastly more than articulate this theory. In Chapter 3 we will return to his account of linguistic meaning as well as to the parallel architecture of the language faculty. I will eventually argue that Jackendoff's work plays an important role in the solution to Benacerraf's dilemma.

[53] See Margolis and Laurence [1999] for useful a discussion. The contrast, of course, is with Fregean Concepts: mind-independent, platonic forms. I address platonism in the next chapter.

[54] Let me mention that while there is broad agreement among cognitive researchers that the integrative work I am describing does indeed take place, not everyone agrees that it is performed by amodal representations. See especially Barsalou [1999].

abstract entities by learning what sort of inferences they support. This occurs via a process of acculturation and social tuning [Jackendoff 2002].

Mathematical entities are paradigmantic instances of abstract entities. If the conceptualist account is to work, it must work for them. Jackendoff rejects the suggestion that the identity conditions of conceptual representations depend on their intentional links with their external, worldly referents. In any case, since many mathematical theories are developed with no canonical empirical application, allowing such links would be of debatable value. The alternative is that the essential features of such representations are exhausted by their inferential links to other representations (their location in the web of belief, if you prefer). What is unclear, on such an account, is how to make sense of a robust notion of mathematical truth distinct from a prevailing community consensus regarding what inferences ought to be drawn. Consider the situation of non-Euclidean geometries before Gauss. Essentially all mathematicians inferred from the information at hand that such geometries were impossible (if indeed they so much as entertained the idea). We now know that they were wrong; they were making a mistake. But to say that requires a standard of correctness distinct from the inferences the community in fact made; indeed, it requires something even stronger than the set of all inferential dispositions the community of mathematicians then possessed. It seems that without a discussion of mathematical truth, conceptualism cannot easily explain how it avoids the slide toward conventionalism.[55]

Unlike the conventionalist, the NORM realist proposes that apart from inferential links, mathematical concepts possess an objective component. Some realists construe this as a form of perception. Here, for example, is Alain Connes:

> The mathematician develops a special sense, I think—irreducible to sight, hearing, or touch—that allows her to perceive a reality every bit as constraining as physical reality, but one that's far more stable than physical reality. . . Exploring the geography of mathematics, little by little the mathematician perceives the contours and structure of an incredibly rich world. Gradually she develops a sensitivity to the notion of simplicity that opens up access to new, wholly unsuspected regions of the mathematical landscape. [Changeaux & Connes 1995]

The challenge for any mathematical realist is to develop an account of how this mathematical perception might work. If we follow Connes in supposing that there exists an objective mathematical reality into which the mathematician develops a peculiar insight then mathematical error becomes no harder to explain than optical illusions. Similarly, the resistance that mathematical reality offers our representations can be attributed to its inherent, mind-independed structure. Finally, the amplicative applicability of mathematics to the study of the empirical world is due, according to the realist, to a pre-established harmony between

[55]Sam Scott and I make this point in Jerzykiewicz and Scott [2003].

mathematical reality (whatever it proves to be) and the structure of empirical reality. None of these answers are available to a conventionalist. In fact, it seems that conventionalism offers little hope of explaining these features of mathematical judgements.[56] We are better off, I think, accepting NORM.

Conclusion

Recall that naked ontological realism about mathematics comprises two theses:

($\mathcal{R}$.i) Some mathematical entities exist; and
($\mathcal{R}$.ii) Their existence is independent of human minds, cultures, languages, and conventions.

I have argued that both of these theses are well-motivated. We have seen that mathematical judgements cannot plausibly be construed as content-free or false. The ampliative indispensability of mathematics in extending the scope of our scientific knowledge discounts the former possibility; while the utility of mathematics in scientific deduction makes the latter implausible. We have good reason to recognize therefore that some mathematical judgements are true. And this, together with the semantic theory discussed in this chapter leads us provisionally to recognize the reality of (at least some) mathematical entities. Moreover, the conduct of standard mathematics gives us a *prima facie* reason to suppose that at least some mathematical entities enjoy an objective, independent existence—though precisely what that amounts to remains obscure. In any event, NORM looks like a promising point of departure. Of course, it's not possible to rest here since many interesting questions remain unaddressed. Among them is the fundamental epistemic problem: how can our cognitive access to the contents of mathematical judgements be explained? Let us now turn to platonism, a popular elaboration of the sort of basic realism I have been defending, to see how this issue might be addressed.

[56] I will use the term 'judgement' to mean either the intentional object or the act of judging. Where possible, I will avoid using the term 'proposition' since it now seems nearly universally to be understood to denote an abstract object.

2 Troubles for Platonism

There can no more be a species of naturalism
that is consistent with belief in the existence of abstract objects
than there can be a species of atheism
that is consistent with belief in the existence of God.
— *Jerrold Katz*

Let's agree on the strength of the evidence just presented that naked ontological realism (or NORM) is on the right track. As I have emphasized, NORM is merely a rough preliminary sketch of a philosophical position. We are still some distance from a fully developed theory. In this chapter I examine how one influential species of abstract realism—Stewart Shapiro's [1997] *ante rem* structuralism—elaborates and extends the basic realist framework.

Although I will focus on one particular account, the point I want to make is general. The vast majority of realists today, including Shapiro himself, hold that mathematical entities (or structures) are abstract and acausal. 'Realism' has, in fact, come to be nearly synonymous with 'platonism'. I think it's important for philosophically-minded cognitive scientists to recognize that this is an error. Many of us today share a malaise concerning abstract realism's apparent inability to explain our knowledge of its posits. In this chapter, I will argue that Shapiro's own account cannot be accepted as overcoming these difficulties. We shall see, moreover, that some of the difficulties Shapiro encounters are symptomatic of the shortcomings of platonist theories as such. If the acceptance of *ante rem* posits lands the realist in trouble then, I want to suggest, mathematical realists are well advised to cut their losses, return to NORM and try something new.

Abstract Realism

Shapiro [1997] characterizes philosophical realism concerning mathematics as the attempt to take the discourse of working mathematicians at face value.[57] To construe realism in this manner is already to take a nontrivial step beyond minimal ontological realism. As we saw in the last chapter, NORM attempts to treat *the intersection* of mathematicians' claims as ontologically committing. It is therefore conciliatory toward constructivism in a way that Shapiro's realism is not. Nonetheless, I propose to concede for the sake of discussion in this chapter that the appropriate point of departure is NORM augmented by whatever Shapiro's classical mathematician wishes to add. Having noted the concession, let us now retrace the chain of reasoning that leads from the arguments of the previous chapter to modern platonism.

[57] Additional examples of this approach can be found in the work of Balaguer [1998] and of Resnik [1997].

Platonism

Mathematics studies mathematical objects: sets, functions, groups, numbers, and so on. It characterizes the essential, formal properties of those objects. But what, apart from their naked existence and formal character, can we assert of such entities? On this question, the mathematician herself remains silent. And so, the first hurdle confronting a philosopher interested in mathematical objects' nonformal properties concerns methodology: by what means can she determine the answer to this question? One apparently promising way forward is to proceed by conceptual analysis. We can start by drawing up a list of properties (or property types) that mathematicalia *cannot* exhibit on pain of contradiction or incoherence, thereby arriving at a sort of photonegative of the solution.[58]

Here's an obvious start: Mathematical entities are not observable—at least not in any straightforward sense. They can be represented by means of diagrams, numerals and other such aids. But even so, mathematical proofs are not *about* the diagrams or the notation; they are about what those represent. Moreover, the unobservability enjoyed by mathematicalia is of a peculiar sort. To see why, let me draw a rough and ready distinction here between two kinds of unobservables. On the one hand, we have entities undetectable by the unaided sensorium in virtue of sheer size, velocity, intense gravitational pull, or some other physical feature. Among such entities we find viruses, quarks, tectonic plates, black holes, and dark matter. A distinct category comprises entities which cannot be touched or seen in virtue of being complexes realized in (or supervening on) other material entities. Instances here include computer files, immune systems, and language acquisition devices. I am not suggesting that this bifurcation is either exclusive or exhaustive. Still, the point can be made that mathematical entities do not fit well on either side. Unlike the cosmologist or geophysicist, the mathematician does not devise or build complex devices so as to better observe her chosen objects. So the obstacle to observing mathematicalia does not seem to stem from our imperfect sensorium [Brown 1999]. Nor does it seem plausible that mathematical entities could be unobservable in virtue of being complexes instantiated in more simple physical systems. One important reason, noted by Frege [1953], has to do with the cardinality of physical entities. Consider: all physical entities are located in space and time. One standard way to construe space-time is as a set of (at most) $2^{\aleph_0}$ points. There are, moreover, at most, a finite number of physical entities in each region of space-time. If so, then the number of physical entities is bounded. There are, in fact, no more than $2^{\aleph_0^{\aleph_0}}$ of them [Parsons 1975].[59] Clearly, this is a colossal number. Still, standard set theory (with the axiom of replacement) permits the construction of sets of the cardinality $\aleph_\omega$— the first cardinal preceded by infinitely many cardinals, so one vastly larger than any collection of

[58] This strategy is mentioned, for instance, by Burgess and Rosen [1997].

[59] To be on the safe side, I'm assuming here that *no more than* a countable infinity of distinct entities can occupy each space-time point. This is probably overly generous.

physical objects could posssibly be.[60] And the iterative hierarchy climbs higher still. So although it may not be immediately clear to a philosopher what it means for these dizzying collections to 'exist,' it cannot mean that they exist in virtue of being instantiated in concrete physical models.[61]

There is, additionally, a second reason not to equate mathematical entities with physical objects. I have already noted that mathematical facts are metaphysically necessary; it could not have been the case that they were otherwise. Try as we might, we cannot imagine coherent possible worlds where (say) there exists a largest prime or where $\sqrt{2}$ is a rational fraction. By contrast, all physical facts are (arguably) in principle contingent. As far as we know, even very fundamental physical facts, such as the values of physical constants, might have been otherwise. The possibility, in any case, appears coherent even if possible worlds where the values are very different from the actual would be uninhabitable by us. The necessity of mathematical facts and contingency of physical ones makes it implausible that the former can be equated with the latter.

Continuing with our conceptual analysis, we find another clue to the properties of mathematical entities in ordinary linguistic usage. To ask *where* ω-sequences are located or *for how long* the conic sections have existed is to pose nonsensical questions. The problem is not that we are currently ignorant of the answers. Rather, it's hard to make sufficient sense of what is being asked to know how to go about formulating a reply. These questions commit what Ryle [1949] called a *category mistake*; they attempt to apply a predicate to a subject matter which is inherently unsuitable to it.

Taken together, these considerations permit us now to venture a first, tentative step beyond naked realism. The position arrived at, *modern platonism*, is characterized by the acceptance of two hypotheses over and above what NORM already commits us to. The first is this:

($\mathcal{P}$.i) Mathematicalia are imperceptible, atemporal, and nonspatial.

The second hypothesis follows immediately from the first. As far as we know, causal interactions involve entities (roughly) localizable in space-time. But since mathematicalia seem neither to be spatial nor temporal, it follows that they are incapable of playing a role in causal interactions. They cannot be generated, do not decay, and cannot be destroyed. Needless to say, this makes them highly unusual (and perhaps even unique) objects.

($\mathcal{P}$.ii) Mathematicalia are incapable of entering into causal interactions.

[60] See Maddy [1997], pp. 57-60 for a discussion of this topic. Also see DeLong [2004] and Enderton [1977].

[61] This leaves open the question of whether mathematical entities can be identified with *possible* physical objects. I defer discussion of modal construals of mathematics for another time.

Modern platonism is sometimes called an *ante rem* (or 'before substance') realism since the platonist holds that the existence of mathematicalia is independent of and metaphysically prior to the existence of extended substance (*res extensa*).[62] Next, I'd like to take a closer look at what is perhaps the most sophisticated current version of *ante rem* realism: Stewart Shapiro's [1997] structuralism.

Structuralism

Stewart Shapiro [1997] endorses both ($\mathcal{P}$.i) and ($\mathcal{P}$.ii). But it would be misleading to call his *ante rem* structuralism a species of platonism without further comment.[63] Indeed, it's only fair to briefly clarify where the differences between standard platonism and abstract structuralism lie.

Consider for a moment the sorts of entities that, according to the platonist, populate the world of mathematics. We find the mathematical realm teeming with groups, numbers, graphs, functions, sets, classes and other, more exotic species. It's worth enquiring whether all of these entities are *sui generis* or whether some are ontologically more basic than others. As is well known, the mathematical realist can make significant economies in her basic ontology by supposing that, in the final count, almost all denizens of the platonic realm reduce to sets. That's because (remarkably enough) we can construct surrogates for just about all mathematical objects by using only the primitive notions of membership and the null set.[64] This does not, of course, imply that all mathematics *is just* set theory any more than the ontologically basic status of subatomic particles entails that all physical science is just particle physics. The various branches of mathematics have their distinctive intellectual styles, techniques, and problems. Still, by recognizing the logical priority of sets we gain a natural way of organizing our ontology.

Taking set theory as ontologically fundamental raises certain problems. It's reasonable to suppose, Benacerraf [1965] argues, that *bona fide* entities have stable identity conditions. Suppose we know, for example, that a certain amino acid is really a particular organic molecule. The amino acid's various properties (molecular mass, polarity, acidity or basicity and so on) can be accounted for directly in terms of its underlying chemical structure. Moreover, armed with our chemical analysis, we can sensibly ask whether (for instance) the amino acid in question contains or fails to contain a sulphur group (or whatever). In the case of the proposed identification of complex mathematical entities with sets, things are

[62] It's useful to notice that platonism is a form of ontological dualism. Note also that some of the same argumentative strategies that are normally deployed in favour of mind-body dualism—in particular, arguments from conceivability and from category mistakes—play a role in sustaining platonism.

[63] The details of the structuralist conception of mathematics won't play a significant role in my argument for the moment. But I want to leave that door open.

[64] Set theory is not our only choice here; one can take mappings as basic and avail oneself of category theory instead. Cf. Hellman [2003]. I return to this topic in later chapters.

not nearly as clean. Consider the natural numbers. If numbers are objects, we should be able to specify precisely what kind of objects they are. If they are *au fond* sets then we should be able to say exactly *which* sets. As is well known however, there is an infinite variety of non-equivalent ways of constructing a bijective mapping from sets to the natural numbers. Because of this, apparently straightforward factual questions—such as whether or not $1 \in 3$—cannot be settled except by fiat. (If we identify the positive integers with von Neumann's ordinals then 1 is indeed an element of 3; if we adopt Zermelo's characterization instead, it's not.) Moreover, the problem ramifies. There are many nonequivanent ways of using sets to offer surrogates for the integers, rationals, reals, and so on. But if that's right, Benacerraf argues, this speaks against identifying numbers, or mathematical entities in general, with sets. In fact it speaks against the notion that mathematicalia are *bona fide* entities at all.

Ante rem structuralism offers an elegant reply. The apparent difficulty stems from supposing that mathematical entities are to be thought of as being characterized by their essential, intrinsic properties. There is however an alternative:

> Mathematical objects [so numbers, groups, sets] are featureless, abstract positions in structures (or more suggestively, patterns);. . . paradigm mathematical objects are geometric points, whose identities are fixed only through their relationships to each other. [Resnik 1997]

On the structuralist account, rather than investigating discrete entities with (as it were) a mysterious inner nature, mathematics studies positions in abstract patterns. The nature of a mathematical object is fully determined by the place it occupies in such a pattern—which is to say, by its external relations to the structure's other positions. The picturesque metaphor can be given precise content. Shapiro [1997] offers axioms that detail the nature of abstract structures. These closely parallel second-order ZF axioms, thus ensuring that the structuralist's proposal is sufficiently rich to offer a background ontology for the whole of mathematics (on the assumption, of course, that ZF does). The apparent problem of multiple reductions that Benacerraf [1965] points out emerges as a natural corollary of the theory. Comparisons between elements *within* a single structure (such as questions whether $1 < 3$, and so on) are perfectly sensible and receive answers. But since structuralism only specifies objects 'up to isomorphism,' comparisons between objects *across* different structures find no principled solutions. Nor would we expect them to. To be a natural number *just is* to play a role in a structure specified by (second order, so categorical) arithmetic. It should not surprise us that a variety of distinct sets are able to play this role.[65]

[65]Note also that structuralism starts us on the path to explaining why the study of mathematics is so useful in natural science: mathematical patterns can serve as descriptions of concrete, physical systems when the latter happen to display a structure isomorphic to the former. I return to this in Chapter 4.

For our purposes, the important point to take away is that abstract structuralism takes on board both of the propositions that more traditional platonism is committed to. In addition, it accepts a further, logically independent postulate:

($\mathcal{P}$.iii) Mathematicalia are fully defined by their formal relational properties.

This postulate is logically independent of ($\mathcal{P}$.i) and ($\mathcal{P}$.ii). It stands or falls independently of them.

Knowledge

Earlier, I mentioned the pervasive perception that mathematical realists (of all persuasions) face some tough questions about epistemology. Addressing this topic, Benacerraf [1973] argues as follows: In order to come to know a new fact F several conditions must be satisfied. Most obviously, F must actually hold. As well, since knowledge is a species of belief, coming to know a fact involves either forming a new belief or shifting preexisting and erroneous beliefs with regard to F. Finally, to count as knowing that F, as opposed to merely having made a lucky guess, we require that appropriate evidentiary grounds exist for our epistemic state. Those grounds are typically construed as the existence of an appropriate connection between ourselves and that which is known. In the case of our knowledge of perceptually observable physical phenomena this connection can plausibly be traced to the causal interaction between our sensory apparatus and our surroundings. Where no connection exists, or in cases where the existing connection fails to be sensitive to the appropriate facts, we cannot be said to have knowledge of F.[66]

What troubles Benacerraf about our apparent knowledge of mathematical entities (as construed by platonists and their intellectual successors) is that it's very hard to say what the relevant epistemic grounds might be. It's implausible that our knowledge of mathematics beyond grade-school—the sort needed to grasp ZFC, for instance—is innately given. Mathematical research is just too difficult and time-consuming for wholesale nativism to be a plausible hypothesis. Yet since abstract entities (including *ante rem* structures) do not, *ex hypothesi*, enter into causal interactions, we cannot easily explain our knowledge of them by analogy with our knowledge of physical facts. Admittedly, it might seem tempting to pass the buck to the mathematician by suggesting that accepted mathematical proofs themselves constitute sufficient evidence for the existence of the relevant objects. Proofs, after all, are considered sufficient evidence for the existence of the relevant mathematical structures by working mathematicians. But, Benacerraf argues, this should not satisfy an

[66]I will discuss knowledge in more detail in Chapter 4. I should say however that I will be focusing here on knowledge construed as true, reliable belief. In other words, I am interested in what makes mathematical judgements true and how it comes to pass that such judgements are often reliable. If there is more to knowledge than that, I am leaving it out of the discussion.

ontologist. What is at issue for her is truth and not theoremhood or "mathematical correctness". Lacking an independent philosophical account of how mathematical theoremhood manages to track mathematical fact one can sensibly deny that we have reliable mathematical knowledge: the proofs might be correct, but the propositions they demonstrate might still not be true.

Benacerraf's argument is not a refutation of platonism. The platonist can reject some of the presuppositions that Benacerraf relies on. Among the more vulnerable premises is the causal account of epistemic grounding [Goldman 1967, Skyrms 1967]. Causal accounts promise an attractive way of circumventing Gettier [1963] scenarios. They do however beg the question against the platonist by building in precisely what she explicitly denies—namely, that knowledge requires causal traffic between knower and known. Of course, the accusation of circularity alone does not prove that causal accounts of epistemic grounding are false.[67] For all we know, some version of such an account may be exactly right. But causal accounts of epistemic grounding cannot be used to construct compelling arguments against the platonist position. On balance then, Benacerraf [1973] should be read as merely pointing out that no generally acceptable account of our mathematical knowledge yet exists.

The burden of proof, I think, distributes evenly to both sides: It's incumbent on skeptics about acausal abstracta to show that a non-circular version of Benacerraf's arguments can be formulated or to furnish an alternative theory which avoids the problematic commitments. The abstract realist, by contrast, needs to make every effort to explain how we might make sense of our knowledge of acausal, abstract facts. Shapiro [1997] takes up this challenge and it is to his account of epistemic contact that I now turn.

Ante rem **account**

Shapiro [1997] proposes an admirably lucid and detailed epistemic account. On that account, we derive our mathematical knowledge from three sources: pattern recognition, linguistic abstraction, and functional (or implicit) definition. Let me briefly explain each of these routes before critically evaluating the hypothesis being advanced.

Pattern recognition. Of the three sources of mathematical knowledge, pattern recognition is the simplest but also the most limited. Typically, the sensorium of an animate creature is capable of registering and distinguishing a certain range of visual, haptic, olfactory and auditory stimuli. The recurring properties of these stimuli—including, for instance, shape, texture, and pitch—can be registered as well. Sufficiently complex creatures are able to track groups or sequences of recurring properties.[68] Human infants are no exception. For

[67]This is sometimes forgotten. See Nozick [1981] for some interesting comments.

[68]For a discussion of some unexpected limitations of this ability in higher mammals, see Spelke [2002].

instance, newborns are already capable of distinguishing patterns consisting of two auditory stimuli from those consisting of three [Bijeljac-Babic et al. 1993]. They are capable of distinguishing faces (especially mother's faces) from other types of patterns and they show a preference for looking at the former [Pascalis et al. 1995]. And there is good evidence that even prior to acquisition of language, infants factor the world in terms of predicates and arguments [Gleitman and Fisher 2005]. Finally, in addition to any innate capacities for specific pattern recognition and stimulus parsing, children are able to learn which *new* types of patterns they need to track; they are, in effect, capable of genuine conceptual advances.[69]

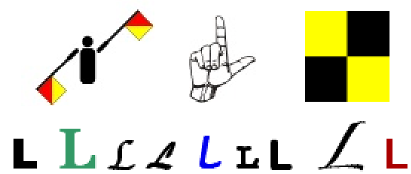


Figure 1: A dozen tokens of a single type.

The *ante rem* structuralist invites us to consider an example of this sort of general learning. Take as our example a child being taught to recognize the letter 'L'. She may start by learning the alphabet song. Part of memorizing the song involves learning that its twelfth term is a particular voiced, alveolar consonant. Next, the child might be taught to correlate occurrences of that phoneme with written tokens of a specific shape. This is not entirely straightforward since written L-tokens vary considerably in visual appearance. Once she masters that skill, she moves on to still more complex challenges. It turns out that the letter 'L' can be tapped in Morse code, gestured in semaphore, and presented as a mariner's flag. What's more, certain letter tokens belonging to non-Roman or non-standard alphabets—including λ, Λ, and Ł—can count under certain conditions as borderline members of the L-type. But in any case, the *ante rem* realist argues that normal children *do* eventually come to recognize context-sensitive, multi-modal patterns and that this occurs on the basis of an exposure to concrete tokens. To make sense of this fact the structuralist suggests that we need to admit that an acquaintance with concrete physical tokens can give rise to knowledge of *places in a pattern.* After all, the only feature shared by the various 'Ls' is the role they play in an alphabet. And an alphabet just is a kind of structure. Thus, to make sense of what takes place, we must come to recognize the existence of structures in addition to individual, concrete objects and their properties.

On the *ante rem* structuralist's reconstruction, something analogous if somewhat more sophisticated takes place when children come to understand natural numbers. Relatively early on, infants recognize that many otherwise very different physical stimuli—such as

[69]Shapiro [2000] suggests that "pattern recognition is a deep and challenging problem in cognitive psychology, and [that] there is no accepted account of the underlying mechanisms." This is not false but it is slightly misleading. Our knowledge of how humans categorize is, of course, incomplete. Nonetheless, standard cognitive science textbooks, including O'Reilly and Munakata [2000] and Gurney [1997], contain fairly sophisticated models and discussions.

light flashes, sequences of tones, and concrete objects—can share a single, higher-order feature: what we would call 'numerosity'. To recognize this, the infant mustn't focus unduly on the accidental physical features of each stimulus but rather attend to multi-modal, context-sensitive information. Adults may perhaps help by actively pointing out various collections while labelling them with the appropriate number term, but in all likelihood the underlying ability is innate [Gallistel et al. 2005]. The important point is that learning to attend to and distinguish collections of various numerosities is analogous (the argument runs) to grasping the alphabet. In the one case, the child learns that certain tokens count as 'Ls'; in the other, she learns that certain collections count as pairs, or fours, or sevens. And this is a first step toward grasping mathematical structures:

> The process. . . may not go all the way to characters and strings as completely freestanding abstract objects, but the development goes pretty far in that direction. Presumably, nothing philosophically occult or scientifically disrespectable has been invoked along the way. In the end, we either demystify numbers [and abstract structures] or we mystify more mundane items [such as letters of the alphabet]. [Shapiro 1997]

In effect, the structuralist argues that in order to explain pattern recognition of the sort employed by young children, we should acknowledge the existence of structures, including the alphabet and the number seven.

Having taken that last step, we appear to face a dilemma: either we insist on thinking of patterns as coextensive with (but not identical to) the elements that comprise them, or we think of them as freestanding and abstract. On reflection, the first option quickly leads to absurd conclusions. For instance, if all tokens of the letter 'L' were destroyed and if the letter truly were coextensive with its token instantiations, then the letter itself would perish. Likewise, if no physical collection of some particular cardinality happened to exist at a given moment, the natural number corresponding to that cardinality would itself (perhaps temporarily?) cease to exist. Recall that we already have reason to believe that there are some infinite numbers that are *never* instantiated in concrete collections. Construing patterns as coextensive with the systems they organize is therefore unacceptable. And so, the structuralist invites us to accept the existence of freestanding, abstract patterns.

Linguistic abstraction. Pattern recognition has serious limitations as a means of acquiring mathematical knowledge. In order to grasp a structure in this manner, one must perceive a concrete system which exemplifies it. The maximum numerosity of systems human beings are capable of perceptually distinguishing is an open question. Shapiro suggests however that it certainly does not exceed ten thousand:

> At some point, still early in our child's education, she develops an ability to understand cardinal and ordinal structures beyond those that she can recognize all at once via pattern recognition and beyond those that she has actually counted or could count.

> What of the 9422 pattern. . . ? Surely, we do not learn about and teach such patterns by simple abstraction and ostensive definition. The parent does not say, “Look over there, that is 9422” [Shapiro 1997, p.117]

The suggestion is that another, more sophisticated process must be invoked to explain our knowledge of more complex structures.[70]

Here, language plays an important facilitating role. As I have noted, we begin to understand the meanings of numerals and number-names once we grasp the connection each of these has with appropriately sized collections. At first, this knowledge is gained piecemeal. But having grasped the connection in the case of small collections, we are ready to take the next step: that is, to realize that numerosity patterns *themselves* form a system. The distinct and systematic labels that language makes available help the child grasp each number as itself an *object*, rather than as a property of collections. Language moreover helps the child understand that the system of numbers (now construed as objects) itself displays a regular, orderly pattern—a pattern with a further, higher-order property: each of its elements has a unique successor, such that no two elements share a successor. Once the child has grasped this, she has (implicitly) grasped the Peano axioms.

Two pieces of evidence speak in favour of this reconstruction. At a certain moment, children delight in making up labels for absurdly large and sometimes nonsensical numbers (“a billion trillion zillion”) and gleefully naming the next higher cardinal. What they seem to be enjoying is their new-found grasp of structures whose corresponding concrete collections they could not possibly envisage. It’s interesting to note moreover that chimpanzees too can be taught to match labels (including arabic numerals) to collections of items with the appropriate numerosity. Interestingly however, chimps take a roughly equal period of time to learn each label for collections from 0 through 9 [Kawai and Matsuzawa 2000]. Unlike human children, they never seem to ‘get’ that every label’s referent must be followed by a successor. One can reasonably hypothesize that our capacity for learning arithmetic somehow piggybacks on our grasp of natural language [Hauser et al. 2002].

Implicit definition. Once learners have at their disposal the full semantic resources made available by natural language, it’s possible to communicate the nature of a structure by indirect description. In the case of structures that make no reference to entities outside of themselves, we can describe the elements that comprise them strictly in terms of the relationships they bear to one another. The system thus described need not have been observed or even to have physically existed. We begin by holding true a plausible collection of propositions or axioms in which some undefined term T appears. Perhaps the collection

[70] There is a second, independent reason why pattern recognition cannot be the whole story: pattern recognition ties mathematical knowledge to sensory experience, while mathematics is typically held up as an example of the *a priori*.

strikes us as self-evident. The term T comes to be implicitly (or functionally) defined by the axioms; it comes to possess whatever meaning it needs to for the statements to come out true. Shapiro suggests that the strategy succeeds provided that two conditions are met. The first is that the sentences that serve to specify T must be consistent; that is, they must be capable of being true simultaneously. Second, the structure specified by the axioms must be unique. Any systems that the definitions hold good of must share a common structure.[71]

Counter-evidence

Shapiro's [1997] account of mathematical knowledge is admirably clear and precise in its commitments. Thanks to its clarity, it is possible to critically examine some of its details. I will start by reevaluating whether linguistic abstraction and implicit definition can indeed account for human knowledge of large structures. I will return to pattern recognition and smaller, finite structures a little later on.

Large structures

According to Shapiro, human beings capable of recognizing relatively small patterns are brought to understand large (and sometimes truly vast) structures by linguistic abstraction and implicit definition. It's worth underscoring that both of these abilities rely crucially on natural language. The theory is quite explicit on this point:

> [These] epistemic techniques suggest a tight link between grasp of language and knowledge of structures. This is especially true for implicit definition. For the fields of pure mathematics at least, grasping a structure and understanding the language of its theory amount to the same thing. There is no more to understanding a structure and having the ability to refer to its places than having an ability to use the language correctly...
>
> [T]he way humans apprehend structures and the way we "divide" the mathematical universe into structures, systems, and objects depends on our linguistic resources. Through successful language use, we structure the objective subject matter. Thus, language provides our epistemic access to mathematical structures. [Shapiro 1997, p.137]

Let me concede for the moment that natural language *could*, in principle, play the required mediating role between acausal structures and our cognitive apparatus. This is a point to which I will return. For now, I want to focus on some of the implications of the proposal. If Shapiro is right—if our epistemic access to mathematical facts is mediated by language—then we would expect severe linguistic impairments to have a deleterious impact on our mathematical abilities. And conversely, we might perhaps also expect impairments of mathematical intelligence to correlate to a degree with impaired linguistic ability (though

[71] A useful introductory presentation of implicit definition can be found in Chapter 2 of Nagel and Newman [2001]. For a rather critical philosophical discussion of the powers of implicit definition, see Horwich [1997].

this is far less certain). A look at the psychological literature does not bear out either of these predictions.

Semantic dementia. The first piece of evidence I'd like to look at comes from studies of dementia. In general, dementias are characterized by a chronic decline in cognitive function across two or more distinct domains. They normally onset gradually and their early symptoms can be relatively mild. The precise clinical profile of affected patients is hard to predict in advance due to the considerable range of underlying neuropathologies. Still, over time dementias are typically more debilitating than those impairments—such as aphasias, ataxias or amnesias—which target only one type of function (so language, coordination, and episodic memory, respectively) [Albert et al. 1999].

My interest here is in semantic dementia.[72] The condition involves a gradual degeneration of semantic memory, typically due to the atrophy of the cerebral left temporal lobe and supporting tissue (Figure 2). Semantic dementia leads to a loss of understanding of the meanings of both spoken and written words, as well as severe difficulties in articulating content. It also results in an inability to recognize objects, faces and pictures. Since other cerebral regions are typically spared, these deficits are circumscribed and most other aspects of mental life remain unaffected. Patients are typically alert and orient normally in their surroundings. Their perceptual faculties, autobiographical (episodic) memory, and problem solving skills remain intact. In fact, patients can even retain the non-semantic (so syntactic and phonological) aspects of their linguistic competence.

Cappelletti et al. [2001] investigate the extent to which a patient affected by semantic dementia retains an understanding of specifically arithmetic concepts and operations. The subject, IH—a 65-year-old, male, right-handed, British banker—was first diagnosed in 1995. IH's initial symptoms included severe difficulties with finding words and also with naming objects. IH remained fluent and his speech was syntactically correct but the investigators characterized his replies as 'vague' and 'discursive', often lacking a clear meaning. To compensate, IH frequently relied on set phrases such as "I delved into that…", or "I am totally committed to…". He was also incapable of reading newspapers. It's important to emphasize however that IH's semantic difficulties were not due to a lack of general intelligence. His episodic memory was largely spared and he continued to recognize people and places. His knowledge of familiar topics, such as sports and politics, also remained intact. The problems IH was experiencing seemed therefore almost wholly connected to his knowledge of language. (There was one notable exception, however: several years earlier IH had begun to display a lack of judgement while gambling. This ultimately led to financial difficulties and divorce. It's not clear whether these symptoms were due to temporal lobe atrophy or perhaps to damage to underlying, subcortical structures.)

[72]This has sometimes also been called progressive aphasia.

The initial investigation comprised a battery of tests designed to measure IH's non-mathematical, semantic knowledge. The stimuli included images of both living and human-made entities drawn from six categories: vegetables, body parts, and animals, as well as furniture, modes of transportation, and musical instruments. In each case, the stimuli ranged from the very typical (tomato, dog) to the atypical (asparagus, zebra). Both verbal and pictorial tests were carried out. The verbal tests asked IH to perform such tasks as picture naming, picture matching by type, producing words semantically associated with a given stimulus, offering verbal definitions of objects, and naming as many objects of a particular type as possible. The results were almost uniformly discouraging. IH was wholly unable to perform these tasks, scoring zero on all but name-to-picture matching. There, since the response options were limited, he performed at chance.[73] Age-matched control subjects, by contrast, performed at 97% or better on all tasks.[74] We must therefore conclude that IH's verbal semantic knowledge was all but nonexistent. IH did somewhat better on pictorial tests of semantic knowledge. Here, the tasks involved understanding and manipulation of pictures chosen so as to parallel stimuli used in the previous tasks. IH was asked to classify pictures both at the entry-level category (animal, furniture) and at the subcategory level (exotic animals). He was also asked to draw size comparisons between depicted items, perform semantic picture-picture associations, and distinguish between real and nonsense objects based on silhouettes. In each case, the responses IH was asked to make were nonverbal to block the interpretation that his deficits had to do with linguistic articulation. He scored 80% on the picture classification at the category level and 66% at the subcategory level. Similarly, he scored 65% on the size-judgement task, and 70% on the silhouette reading task. (Predictably, the controls' scores were nearly perfect for all but the silhouette task.) The results suggest that while IH's linguistic semantic abilities were almost wholly compromised, he did retain some ability to understand objects and their properties.

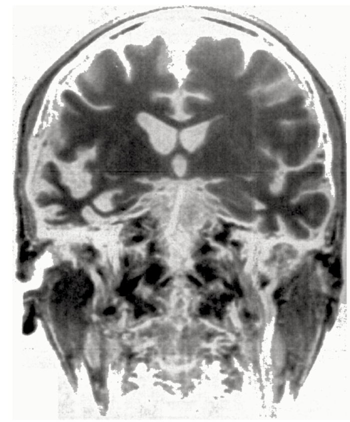


Figure 2: Left temporal cortical atrophy resulting in semantic dementia [Cappelletti et al. 2001]. Used with permission.

[73]The sole exception to this pattern of results concerned the naming of seasons, days of the week, and months of the year where IH scored 21 out of a possible 23. This, together with the rest of the results, raises the question whether rote patterns such as the list of the days of the week are encoded differently from other semantic memories. A difference in encoding is further suggested by neurophysiological studies, including Dehaene et al. [2003].

[74]In the category fluency test, IH named zero objects of the types used; the controls averaged almost 15 per minute across the categories.

The principal purpose of the study was, of course, to investigate IH's specifically *mathematical* abilities. If these abilities depend on our capacity to grasp linguistic meaning, one would expect IH's understanding of mathematical concepts and operations to be very poor. If, on the other hand, the two are importantly distinct, at least some mathematical understanding should be spared in spite of IH's other semantic difficulties. Two broad types of tasks were used: the first focused on comprehension and transcoding, the second on calculation. In the first set of tests, IH performed nearly flawlessly. The tasks included recognizing written numbers, counting, naming a number's successor and predecessor, transcoding from arabic numerals to spoken number words and vice-versa. He made one mistake (in ten trials) when asked to bisect numbers. The only exception to his apparently nearly perfect comprehension of numbers involved knowledge of number facts. He was not able to say how old he was, what his shoe size was, or how many months there were in the year. Moreover, interestingly, he was not able to name or explain the arithmetical the operators. Still, IH clearly retained much of his understanding of numbers and their properties.

Given IH's apparent difficulties with explaining the arithmetic operators, one might expect him to have trouble with written calculation. Not so. In fact, he scored above 95% on 2 and 3 digit addition and subtraction problems and 69% and 62% on multiplication and division problems, respectively. His single-digit arithmetic performance was even better. It seems therefore that his knowledge of these operations is largely preserved in spite of the clear semantic impairments he displays. (Cf. Table 1, below.)

Task	*IH % correct*	*Controls % correct*
Oral single-digit arithmetic (N=254)		
Addition	98	100
Subtraction	95	100
Multiplication	73	90
Written multidigit arithmetic (N=128)		
Addition	99	99
Subtraction	96	98
Multiplication	69	95
Division	62	95
Approximate calculation		
Approximation to correct result	not understood	100
Placing numbers on a line (N=100)	100	100

Table 1: Semantic dementia impacts arithmetic cognition (cf. Cappelletti [2001]).

We should conclude therefore that while IH has very serious nonverbal semantic memory deficits, his understanding of numbers and of arithmetic procedures remains largely intact. This strongly suggests that our understanding of mathematical and of linguistic items involve independent cognitive systems. It therefore gives us our first evidence that mathematical understanding operates independently of some aspects of our linguistic competence.

Agrammatic aphasia. The case of IH leaves open the possibility that our mathematical abilities are bound up closely with syntactic (rather than semantic) aspects of our linguistic competence. Evidence showing that this is not the case comes from work on agrammatic aphasia.

In what (to my knowledge) is the first study of its kind, Varley et al. [2005] investigated the arithmetic abilities of three profoundly aphasic men: S.A., S.O., and P.R.[75] All three subjects were in their late 50s. All three had suffered lesions to their left middle cerebral artery resulting in extensive damage to the left perisylvian temporal, parietal and frontal cortices. Consistent with this damage, the subjects displayed severe but circumscribed linguistic deficits. All three performed above 85% on both spoken and written word-picture matching tasks. And two of the three did relatively well ($> 75\%$) on spoken and written synonym matching tasks.[76] In each case, the subjects' phonological memory was also relatively spared. Nevertheless, the subjects performed poorly on tests of grammatical processing involving the matching of reversible spoken and written sentences to pictures depicting relevant actions (for example "The man killed the lion" and "The lion killed the man"), scoring below chance on this task. Since their word-knowledge and linguistic memory were apparently not a factor, their failure on this task can only be attributed to a specific grammatical deficit.

Interestingly, all three subjects retained considerable mathematical competence. They were able to add, subtract, multiply and divide whole numbers. They were also able to add and subtract fractions. Moreover, in spite of their difficulties with reversible sentences, the subjects did relatively well with reversible subtraction and division problems. The simplest task of this sort involved solving pairs of arithmetic expressions—such as $59-13$ and $13-59$, $60 \div 12$ and $12 \div 60$. In order to arrive at the answer, the subjects needed to keep track of the order of presentation and understand its impact on the calculation being performed. A second, slightly harder task of a similar nature used bracketed expressions $(36 \div (3 \times 2))$ that the subjects were asked to solve. Once again, the subjects performed relatively well. The third and hardest task had the subjects insert brackets into unbracketed expressions (such as $7 + 4 \times 3 + 17$). Here, the subjects were deemed to have succeeded on a trial if they were able to insert the brackets in two distinct ways into the given expression so as

[75] Also relevant is a commentary by Brannon [2005] and the work of Gelman and Butterworth [2005].

[76] The remaining subject scored 70% on the written part of this test.

to produce two different answers. The arithmetic task results are summarized in Table 2, below.

Task	*S.A*	*S.O*	*P.R.*
Calculation tests (N=20)			
Addition	19	16	20
Subtraction	19	19	19
Multiplication	19	13	17
Division	19	11	16
Adding and subtracting fractions (N=30)	27	27	20
Reversibility tests (N=40)			
Subtraction	40	35	37
Division	37	34	38
Bracket expressions			
Calculation accuracy (N=64)	45	52	43
Bracket generation and calculation (N=5)	4	4	2

Table 2: Agrammatic aphasia impacts arithmetic cognition (cf. Varley [2005]).

On the basis of the above results, we must conclude that all three test subjects were capable of coping with complex hierarchical structures in the context of arithmetic operations. It seems that they could understand and apply hierarchical, syntactic reasoning regarding arithmetic problems that they were unable to bring to bear on linguistic expressions. This suggests that mathematical processing can operate independently of the processing of natural language syntax.

Savants. Work presented thus far suggests a certain degree of independence between human numerical and linguistic abilities. Nonetheless, it doesn't yet pose an insurmountable problem for accounts of mathematical knowledge such as Shapiro's. An *ante rem* structuralist can maintain that a grasp of natural language (syntax or semantics) is required in order to *initially learn* mathematical concepts and operations; once these are understood, however, our mathematical abilities operate (or degrade) independently of language.[77] My

[77] *Ante rem* structuralists are not the only ones who advance this hypothesis. The dependence of arithmetic competence on natural language is also defended by Hauser et al. [2002]; limited empirical support is offered by Donlan et al. [2007].

next piece of evidence is intended to block this move. It concerns the highly unusual skills of an autistic savant calculator.

Mathematical savants are able to appreciate relations between numbers not apparent to the rest of us. The mathematician G. H. Hardy (whom we encountered in Chapter 1) reports a conversation with Ramanujan, a mathematical genius, while the latter lay dying of tuberculosis in a British sanatorium.

> "The taxi that I hired to come here bore the number 1729," said Hardy. "It seemed a rather dull number," "Oh no, Hardy" replied Ramanujan. "It's a captivating one: It's the smallest number that can be expressed in two different ways as a sum of two cubes"—$1\ 729 = 1^3 + 12^3 = 10^3 + 9^3$. [Quoted in Dehaene 1997, p.148]

The intuitive familiarity with natural numbers required to make this sort of observation is a rare gift that is far from being understood by modern cognitive science. The ability reportedly correlates (in many cases, at least) with a certain sensibility which facilitates mathematical research. Indeed, a number of gifted mathematicians, including Gauss, were calculating prodigies. Nonetheless, Dehaene [1997] argues that at least some apparently superhuman feats of calculation rely on heuristics that can be learned and practised. Moreover, calculation ability alone does not necessarily correlate with general intelligence or the capacity to construct imaginative solutions to novel abstract problems.

The case that interests me here is that of Michael, a young savant calculator. Michael is doubly unusual: not only is he a gifted calculator, he is also profoundly autistic. Autism has become something of a *cause célèbre* in the past two decades. The diagnosis spans a range of disorders whose physiological basis is still not well understood. Autism-spectrum disorders can however be characterized cognitively as involving a characteristic pattern of *executive*, *social*, and *linguistic* deficits. Many autists engage in stereotyped, repetitive behaviours and display obsessive interests. They dislike changes in routine. And they have trouble shifting attention in a flexible and appropriate manner. When focused on a stimulus, they display a bias for local, part-oriented processing. Perhaps for this reason, they seem not to succumb to some visual illusions involving gestalt patterns. Three quarters of autists have an IQ in the mentally retarded range, though some are of average or even above average intelligence. Even high-functioning autists have trouble attributing mental states to others (or to themselves) and so they typically have severe difficulties interpreting or predicting others' behaviour in terms of beliefs or desires. Many autists display serious difficulties with linguistic communication, with special difficulties in the area of pragmatics—they tend to grasp the literal rather than the intended meaning of what is said. They don't get jokes or respond appropriately to metaphor [Frith and Happé 1999, Tager-Flusberg et al. 2001].

Michael presents with the classic symptoms of autism with respect to executive function, social behaviour and linguistic ability. He does relatively poorly on tests of general intelligence (IQ 67). However, he is fascinated by jigsaw puzzles which he solves equally well with the picture-side up or down. He is also interested in calendars. Thus, on tests of intelligence involving only abstract shapes and relations, Michael scores well above average (IQ 140). Michael never initiates social interaction and shows no interest in it. Nor does he have any interest in communicating with others. He does not point and does not attend to pointing. He was once taught some rudimentary signs but never uses these spontaneously. Finally, Michael has never learned to speak and shows no indication of understanding spoken language. In fact, he is, by all accounts, completely alinguistic. This is, I realize, a surprising and radical claim, so I quote the relevant descriptions in full:

> Michael is a young man without any speech or verbal comprehension. . . As a young child, he did not talk or attempt to engage in any kind of communication. He still cannot speak but has learned to copy numbers and letters, though only very poorly. [Anderson et al. 1999, p.385-7]

> He did not look at things when someone pointed at them, never waved goodbye or responded to cuddling. Michael is not deaf, but he seemed not to understand any language at all and did not himself develop any speech. . . He has remained entirely without language, and though he was taught some sign language gestures he never used these spontaneously.[Hermelin 2001, p.109]

> [H]e never initiated gestures, such as pointing or waving goodbye. He never began to speak and did not respond to language. He took very little interest in adults and did not try to communicate in any way. . . He began attending a special school for autistic children at age six. He learned to 'write' with a pencil, i.e., he learned to copy letters and numbers. But he has not improved in this skill since his schooldays and his written numbers are often difficult to make out. He also learned a few elementary Paget Gorman signs, though he never used them spontaneously. [Hermelin and O'Connor 1990, p.165]

In one case, there is some suggestion that Michael's linguistic handicap runs even deeper than the above passages suggest. Anderson et al. [1999] write that not only is Michael alinguistic; he also lacks the underlying resources to categorize pictures of concrete objects into the relevant categories (such as vegetable, mode of transport, and so on). This appears to be indicative of a profound semantic deficit.

> Michael lacks any language production or comprehension. We might go further, and suggest from his performance on the Columbia [inclusion/exclusion test] that he may also be unable to abstract from objects the semantic categories to which they may be assigned. . . [I]t may be that Michael demonstrates above average intelligence only when problems are limited to spatial and perceptual dimensions. Moreover, it seems clear that his capacity to deal with problems as long as they do not involve a semantic classification of objects in the real world also extends to numbers.[Anderson et al. 1999, p.399]

It would appear then that Michael's grasp of natural language (syntax, semantics and pragmatics) is essentially nil.

As it happens, both of Michael's parents are mathematicians. He was taught to 'read' and copy numerals quite early on, though (as we just saw) his writing is often hard to decipher. What is perhaps more surprising given the extent of his various deficits is that Michael is capable of performing basic arithmetic operations on numbers. He can add, subtract, divide, and multiply. Moreover, Michael is capable of factoring numbers. This last ability was the focus of Hermelin and O'Connor's [1990] study.

Three different tasks were used to test the savant's skills: *recognizing* and *generating* primes, as well as *factorizing* non-primes. Each task was performed with three-digit, four-digit and five-digit numbers, so at three levels of difficulty. In each case, the task was modeled for Michael twice. After this he was able to proceed with most of the tasks "appropriately and without hesitation." In one case, a trial had to be rerun to secure Michael's cooperation. Table 3 compares Michael's performance to that of a control subject, a male psychologist with a degree in mathematics.

In general, both subjects' error rates as well as the kinds of mistakes committed were similar. In general, the control tended toward omission errors while Michael tended to produce false positives. The most striking difference between the data concerns the response times. In general, the speed of information processing as measured by reaction time studies is closely associated with the level of general intelligence [Jensen 1979]. In a prior study of idiot-savant calendrical calculators Hermelin and O'Connor [1983] had found that the "simple and complex visual RT of these subjects was in accordance with those expected from their IQ whereas their speed of calendrical calculation was much faster than that usually obtained from people with much higher IQs." This suggests that savant calculators possess a cognitively-specific calculating ability.

It's hard to deny that Michael possesses genuine mathematical knowledge. This is particularly clear from his ability to generate prime numbers in the range between 10037 and 10133. (Recall that Shapiro [1997] specifically denies one could grasp structures as large as 9422 by simple pattern recognition.) Moreover, we can be confident that his knowledge is the same in kind as that displayed by the neurotypical control. An analysis of response times suggests that Michael employs Eratosthenes' sieve, the same algorithm used by the control. The algorithm involves dividing the target number by all prime numbers less than or equal to the target's square root. Thus, for example, since 59 cannot be divided by 2, 3, 5, or 7, we can safely conclude that it is prime. So while Michael's *access* to mathematical knowledge is exceptionally fast, and perhaps unconscious, the knowledge itself appears to be the same in kind as that possessed by the rest of us. It follows that a grasp of language is not strictly required for genuine mathematical knowledge. It would seem therefore that any account of human mathematical knowledge which holds that a knowledge of language is required in order to grasp complex mathematical structures, such as the natural number structure, is empirically disconfirmed.[78]

	Control		*Savant*	
Task	*Correct*	*Mean time (sec)*	*Correct*	*Mean time (sec)*
Recognizing primes				
between 301-393	20/30	11.46	29/30	1.16
1201-1309	18/30	12.90	22/30	2.90
10307-10427	23/30	10.73	15/30	2.00
Generating primes				
227-281	8/10	12.9	9/10	6.20
1019-1091	5/10	25.6	5/10	6.00
10037-10133	4/10	50.0	5/10	10.00
Factorizing numbers				
212-221	8/10	22.6	9/10	8.8
1001-1011	7/10	25.5	8/10	20.8
10002-10013	4/10	48.0	7/10	38.2

Table 3: Savant performance on arithmetic tasks (cf. Hermelin O'Connor [1990]). Note: Results shown indicate the number of correct responses and mean decision time.

Dyscalculia. In order to demonstrate the mutual independence of two cognitive capacities, it's important to show a *double* dissociation: that is, to prove that each can operate (or fail) independently of the other. The last piece of evidence I'd like to present is intended to show that our mathematical competence can fail while leaving the rest of our cognitive capacities intact. Strictly speaking, this is not crucial to the case I am trying to build. But it does serve to reinforce the conclusion that mathematics and language are subserved by functionally independent cognitive systems.[79]

The evidence here comes from studies of dyscalculia, a relatively common but still not well understood developmental disorder [Butterworth 2005]. Dyscalculia seems to affect at least 3.6% of the population—so roughly as many people as dyslexia. Those affected show a persistent impairment learning and remembering arithmetic *facts* as well as problems executing calculating *procedures*. Of course, reasons for poor math skills among children can vary widely. They can include poor teaching, weak study skills, anxiety, missing lessons,

[78]This includes Shapiro [1997] but also Maddy [2007], and arguably the fictionalist accounts of Yablo [2001], and Hoffman [1999].

[79]There is some evidence showing a parallel neurophysiological dissociation. See Kadosh et al. [2007].

and so on. The situation of dyscalculics however is qualitatively different. Their inability to learn is due to persistent problems in representing and retrieving arithmetic information from long-term semantic memory. This manifests as a lack of intuitive knowledge of even the most basic arithmetic. Thus, in spite of normal intelligence, good instruction and concerted effort, dyscalculics can literally fail to understand what a teacher is saying:

> *Child 5*: Oh, there's this really hard thing, about when you're doing times—Ms. S says you can't take away this number, but I keep on taking it away, I don't understand one single bit of it.
> *Child 2*: I sometimes don't understand whatever she (the teacher) says.
> *Child 1*: I don't forget it, I don't even know what she's saying.
> [Butterworth 2005]

Along with problems with grasping the relevant facts, dyscalculics show impairments executing calculation procedures. When they do add, subtract, multiply or divide, they typically do so much more slowly. Their performance is error-prone and they lack confidence in their results. And, even as adults, they rely on immature strategies, such as finger-counting.

Butterworth [1999] presents an interesting case study of dyscalculia. "Charles" is an intelligent and resourceful university graduate in his thirties. He has a degree in psychology and works as a psychological counsellor. As one might expect, he copes well with daily life. However, Charles has had severe difficulties with mathematics since childhood. He cannot add up the price of groceries, count the money in his wallet, or figure out the correct change. When tested, he proved completely unable to solve two-digit subtraction problems. He cannot work out multiplication problems involving numbers greater than 5. And although he can find the solution to single-digit addition and subtraction problems, his performance on these is four times slower than a control subject's (so roughly three seconds). Perhaps Butterworth's most extraordinary findings concern Charles' performance on tasks thought to involve very low-level cognitive abilities. One of these is simple number comparison. In general, the time taken by math-typical subjects to compare two single-digit numbers is (roughly) inversely proportional to the difference between them; it's easier to judge that 2 is smaller than 9 than it is to judge that 8 is. In Charles' case, this pattern is reversed. The time it takes for him to compare two numbers is proportional to the difference between them. This suggests that he is forced to perform number comparison tasks in a way entirely unlike that of typical subjects. This supposition is reinforced by subitizing data. Math-typical subjects take almost the same amount of time to grasp (or 'subitize') the numerosity of collections containing one, two and three items. This capacity is thought to be a very low-level cognitive or perhaps even perceptual ability.[80] Interestingly, Charles does not subitize; he laboriously counts items even in patterns containing two or three entities.

[80] See Trick and Pylyshyn [1993].

This, once again, reinforces the conclusion that dyscalculics are affected by a pervasive, math-specific cognitive impairment.[81]

Discussion. We have now seen some of the evidence pointing to a functional double dissociation between human linguistic and mathematical abilities. In Chapter 4, we will review additional evidence showing that the processing of linguistic and mathematical information is subserved by independent cortical regions. I want to emphasize that the emerging picture is far from straightforward. We need to keep in mind that certain linguistic and mathematical (specifically, arithmetic) abilities *do* appear to interact in various respects. For instance, Welsh and Chinese-speaking children learn the count sequence faster than French or English speakers. This can be attributed to the fact that number-terms in the former two languages are perfectly regular, whereas French and English involve tricky exceptions to a regular pattern ('quatre-vingt', 'eleven') [Miller et al. 2005]. Moreover, studies from Amazonia by Gordon [2004] seem to suggest that users of languages which lack count terms beyond the first three are impaired in their arithmetic abilities. However, in spite of limited interactions, the examples discussed above show that sophisticated mathematical and linguistic capacities can develop and operate independently. And if that's right then any theory which claims that mathematical knowledge piggybacks on linguistic competence is committed to an empirically false picture. We have seen that this is precisely what the *ante rem* structuralist's account of our knowledge of large structures maintains. And so that account must be amended or rejected.

Small structures

As we have seen, Shapiro [1997] himself holds that perceptual pattern recognition does not suffice to explain our ability to track facts about complex mathematical structures—those which comprise more than several hundred places:

> One cannot grasp a structure S by simple pattern recognition unless one can perceive a system that exemplifies S. Such a structure can have at most a small, finite number of places. [p.129]

Still, one could readily imagine an abstract realist adopting a more sanguine stance. Such a theorist might argue—perhaps taking her cue from the cognitive mechanisms discussed by Maddy [1990]—that, when correctly understood, perceptual pattern recognition does

[81] Before moving on, I should mention that (unlike Charles) many dyscalculics discussed in the literature present with a variety of additional cognitive deficits. Some display general working memory problems, left-right discrimination difficulties, spatial and psychomotor deficits, agnosia, dysgraphia, and reading problems. Dyscalculia and dyslexia in particular frequently co-occur. Recent work (for example Landerl et al. [2004]) suggests that while such comorbidity is common, it is dissociable from the core math-specific problems.

indeed afford us epistemic access to even the most arcane mathematical posits.[82] I now want to argue that, on the contrary, pattern recognition (and similar processes) are wholly incapable of offering us an explanation of our knowledge of purported abstract facts.

Let's begin from some shared assumptions. I take it as established that there are mathematical facts and that we have true beliefs about some of them. Of course, we also sometimes make mistakes. But when these are spotted, we respond by changing our minds, tweaking our beliefs, and continuing on with our mathematical research.

Two aspects of this situation deserve to be distinguished and attended to. The first is *aboutness*. Our mathematical beliefs are, of course, about mathematical entities. (What else?) This is easy to ignore since aboutness comes so cheap. No contact of any kind need exist between us and what we think about. In fact, what we think about need not so much as exist. We are just as happy thinking about Boston as Gotham, orchids as phlogiston.[83]

The second aspect of the situation that deserves mention is the *responsiveness* of our beliefs to evidence. We say that a belief is responsive to a set of facts if discovering new information concerning those facts is capable of altering that belief. Typically, when all goes well, our beliefs about real, existing entities are responsive to the states of those entities. Thus, Boston-beliefs are responsive to Boston-facts, orchid-beliefs to orchid-facts, and so forth. This is not the end of the story however. Boston beliefs may additionally be responsive to a host of other facts as well: perhaps facts about baseball, or linguistics, or New England. Sorting out fully and precisely *which* facts a given belief happens to be responsive to is difficult and perhaps even impossible. (Luckily we won't need to do any of that here.) Notice also that the situation is slightly different in the case of beliefs about the non-existent. Our phlogiston-beliefs cannot be responsive to phlogiston-facts; there aren't any such facts. Instead, phlogiston-beliefs are responsive to a variety of other states of affairs—including those involving oxygen, combustion, wood, charcoal, history textbooks, and so on. Again, we may not be able to demarcate precisely which states of affairs are relevant and which are not. But, once again, we must recognize that a fuzzy boundary is a boundary nonetheless; the difficulty of making a sharp distinction does not detract from the overall point.

Consider now what our beliefs *about* mathematical entities are *responsive* to. They cannot be responsive to mathematical facts construed along platonist lines. That's because

[82] Alternatively, an abstract realist might argue that if pattern recognition does not offer us epistemic access to abstracta then some other, relevantly similar procedure does. See, for instance, Resnik [1997]. While I will keep the discussion focused on pattern recognition for the sake of clarity, the overall argumentative strategy employed here extends readily to other such views.

[83] The contrast here, of course, is with reference which, by most accounts, requires at least that the entity referred to exist (though see McGinn [2000]). I mention reference only to set it aside until Chapter 3. The important point for us not to let go of is that our mathematical beliefs are *about* mathematical structures and entities.

the process of altering one's beliefs is a cognitive process. Cognitive processes supervene on neurophysiological events—by which I mean that no change on the cognitive plane occurs without an accompanying neurophysiological change (though not vice-versa). One of the important and remarkable empirical findings of the past century is that all physical processes—including those that take place in living organisms—are fully causally closed. Every physical effect is fully determined by law by antecedent physical occurrences [Papineau 2001, Spurrett 1999]. On the platonic account, mathematical states of affairs are causally inert. No physical process can be altered in response to acausal facts (even if we allow that such facts obtain). Thus, *a fortiori*, no cognitive process can alter in response to abstract mathematical facts. It follows that when we recognize a mistake in our understanding and change our minds so as to have our beliefs accord with mathematical reality, if our beliefs are responsive to something, that something is not the mathematical facts.[84]

What then are mathematical beliefs responsive to? It might be tempting to insist on the noble origin and purity of mathematics by suggesting that mathematical beliefs have no need of facts—and certainly not of physical facts! On reflection however, this proves incoherent. If beliefs about mathematicalia are not responsive to acausal facts *and* they are not responsive to physical facts then they are not responsive to facts, full stop. We have agreed however that we possess mathematical knowledge. Beliefs that are not responsive to (any) facts come in two varieties: they can be irrational *idées fixes* or they can shift utterly randomly. In either case, such beliefs cannot be constitutive of knowledge. Neither the platonist nor anyone else should be driven to characterizing our mathematical beliefs in

[84]The platonist can reply that, in fact, some physical effects *not* determined by antecedent physical processes and hence that my argument does not go through. Radioactive decay has been cited to make this point. I think this counter-argument fails for two reasons.

First, while α-decay is certainly unpredictable, in the sense that no single α-particle emission can be fully predicted in advance, the process as such is lawful. The rate of decay of a radioactive sample at a time t is given by $R = R_0 e^{-\lambda t}$ where R_0 is the decay rate at $t = 0$ and λ is the disintegration constant characteristic of the particular process. In the present context, the point to keep in focus is that the rate of α-decay, and hence the chance of an α-emission, is fully determined by the state of the physical system and nothing else. It's far from obvious how the statistical nature of the underlying physical law can help the platonist's case for the existence of acausal entities.

There is moreover a second reason why an appeal to quantum physics is unhelpful. Suppose one were to concede (for the sake of argument) that it's only at the *macroscopic* level that every physical effect is fully determined by antecedent physical causes. What follows? It has been suggested by Penrose [1989] and by Hameroff [1998] that the explanation for some mental processes lies ultimately with quantum events in neuronal microtubules. Quantum indeterminacy, they argue, plays an important role in our mental lives. The suggestion is indeed intriguing but there is currently *no evidence* suggesting that it's true. Indeed, there is evidence to the contrary: see Churchland [1998] and Franks & Lieb [1998]. The relevance of quantum phenomena to human cognition remains an intruiguing logical possibility. For that possibility to be wrought into an objection in the present context, the platonist would need to show that quantum effects do indeed play a role in human cognition; if they do not, my argument goes through. And she must further explain how such quantum indeterminacy opens the door to acausal objects; for if it doesn't then the discussion is irrelevant.

I am indebted to James Robert Brown for calling my attention these issues.

this way.

There's another option. One might reasonably suppose that only *a posteriori* beliefs are responsive to a circumscribed set of facts; perhaps beliefs about mathematicalia are unusual in being responsive not to this or that state of affairs, but rather to *the totality* of facts. This would perhaps help explain why none of our experiences can disconfirm a true mathematical proposition. There is no doubt a grain of truth here somewhere. Nonetheless, it won't help the platonist. The sum total of causally-potent, physical facts do not add up to an *ante rem*, acausal state of affairs. Even were we to accept this possibility, therefore, we would still be compelled to conclude that our beliefs about mathematicalia are not responsive to what the platonist tells us mathematics is about: the abstract facts of mathematics.

If we allow that our beliefs about mathematics are responsive to some set of facts, but not to mathematical facts, some odd consequences follow. Imagine, for instance, a scenario whereby we (somehow) come to be excommunicated from the platonic realm. Less figuratively: imagine that our beliefs about mathematics continue to be responsive to whatever physical facts they are currently responsive to but that platonic facts (somehow) effectively vanish. What, if anything, would the impact of such an unparallelled metaphysical catastrophe be? A moment's reflection suggests that, in fact, we would not so much as notice. The physical universe would continue to run its course. Our brains would continue to function as they do. And the cognitive processes that working mathematicians undergo would exactly mirror the transitions that they currently undergo. Plausibly, even their phenomenology would remain the same. The same articles would be written. The same proofs would be accepted for publication or rejected. Mathematics would continue to be indispensable to the conduct of natural science. In brief, none of us would be any the wiser. It seems then that mathematicalia construed as platonic abstracta are otiose. They are an unnecessary fifth wheel, an empty place-holder, in explanations of the nature and conduct of the mathematical enterprise. I suppose that one can nonetheless maintain that *ante rem* states of affairs exist. But I'm not sure what the point of this hypothesis that might be. Better, I suggest, to conclude that ($\mathcal{P}$.ii)—the proposition that mathematical states of affairs are acausal—is false.[85]

My argument rests on two vulnerable assumptions. The abstract realist can reject the hypothesis that mental states supervene on brain-states, opting perhaps for some form of mind-body dualism. And she is free to reject the hypothesis that physics is causally complete. In either case, since both of these propositions are today widely held by philosophers and scientists, the burden is on her to prove her case. Lacking such proof, the mathematical realist ought to give up on ($\mathcal{P}$.ii). And since ($\mathcal{P}$.i) entails ($\mathcal{P}$.ii), the rejection of the latter results (by *modus tollens*) in the rejection of the former also. That last step is tantamount

[85]See Rosen [2001] for a related line of argument.

to a wholesale rejection of an *ante rem* conception of mathematics.[86]

Conclusion

The abstract realists' inability to furnish a satisfactory account of our epistemic access to mathematical facts is sometimes read as giving credibility to ontological anti-realism. Hartry Field [1989], for instance, comments:

> Benacerraf's challenge—or, at least, the challenge which his paper suggests to me—is to provide an account of the mechanisms that explain how our beliefs about [abstract] entities can so well reflect the facts about them. The idea is that *if it appears in principle impossible to explain this*, then that tends to *undermine* the belief in mathematical entities, *despite* whatever reason we might have for believing in them. [Field 1989, original emphasis.]

One part of what Field says is, I think, exactly right. Until the *ante rem* realist can offer a plausible epistemic account, we should remain skeptical about her overall conception of the nature of mathematics and its posits. Nonetheless, I think the passage takes a step too far. Abstract realism makes weighty (and questionable) commitments beyond those that an ontological realist must strictly make. The epistemic difficulties in which abstract realists find themselves ensnared are logically independent of the arguments for NORM presented in the previous chapter. A steadfast nominalist may perhaps want to reject the latter as well (though perhaps not on naturalist grounds *per se*). To do that, she would need to offer an alternative explanation of maths' deductive and abductive indispensability.[87] As things stand at the moment, the mathematical realist is free to reject ($\mathcal{P}$.i-ii), to backtrack to NORM and explore alternative possibilities. This is indeed roughly what I propose to do, but not until Chapter 4. Before we can entertain new realist proposals, some of our background assumptions concerning language need to be modified.

[86] None of this entails a rejection of structuralism. Plausibly then, the bulk of Shapiro's [1997] work is very much on the right track in key respects.

[87] Notice that such an account is owed even if Field's nominalization program were to succeed for all of physics, including quantum mechanics. For a discussion of this point, see Steiner [1998].

3 The Semantic Problem

If the semantic and the intentional are real properties of things,
it must be in virtue of their identity with properties
that are neither *intentional* nor *semantic.*
If aboutness is real, it must really be something else.
—Jerry Fodor

Although I did not make much of this earlier, Benacerraf's classic [1973] paper actually raises two distinct problems. Or—to put this another way—it raises one perfectly *general* philosophical problem and quickly moves to discuss one tightly constrained, *narrow* version of it. The general problem, as we have seen, is to find some way to simultaneously succeed at two tasks. We need to furnish a satisfactory theory of meaning for sentences that (realists will agree) express mathematical knowledge, sentences like the one we encountered earlier:

(5) Seventeen is a prime number.

At the same time, we are asked to provide an account of how creatures such as ourselves come to have epistemic access to the contents of those sentences. If it's the objective of philosophers to articulate how 'things, in the broadest sense, hang together in the broadest sense,' then this is surely a worthwhile undertaking.[88] The narrow problem is a whole other matter. In order to make headway on the general problem Benacerraf, quite understandably, adopts what he takes to be plausible assumptions—in particular, he makes rather substantive commitments concerning the nature of human languages, the appropriate way to go about constructing a theory of meaning, and the semantics of mathematical expressions. Regardless of the correctness of those assumptions, it's important to notice that attempting to solve the general problem having built in additional premises constitutes a significantly narrower, more circumscribed enterprise. If moreover the particular assumptions Benacerraf embraces are subtly incorrect in some way then it's possible that the narrower version of the problem has no solution even though the general problem does. In the first part of this chapter I will argue that this is precisely our situation: Benacerraf's conception of natural language—and hence of the 'language' of mathematics—is off the mark in several nontrivial respects. To make headway on the general version of Benacerraf's problem we will therefore need to backtrack and try a fresh approach.

The positive portion of this chapter is concerned with explaining what such a fresh approach should look like. In a nutshell, I think we need to commit to the Chomskyan paradigm both with respect to natural language syntax and with respect to semantics. For reasons that will become apparent in the course of the discussion, this will require a substantial reinterpretation of NORM. In particular, I will argue that we should move from

[88] The characterization, of course, is owed to Sellars [1962].

an *extensional* to an *intensional* realism regarding mathematics; viz., rather than positing objective mathematical *entities*, our minimal ontological realism should commit us to profoundly non-negotiable *constraints* on the formation of correct mathematical judgements.[89] The subsequent chapter will be devoted to laying out an account of how objective constraints of the appropriate type can arise.

Benacerraf's assumptions

What then *are* the assumptions that Benacerraf [1973] relies on to constrain the general problem concerning mathematical meaning and mathematical knowledge? The following passage encapsulates his view nicely. I have inserted markers for the sake of clarity.

> [1] The semantical apparatus of mathematics [should] be seen as part and parcel of that of the natural language *in which it is done*, and thus [2] whatever *semantical* account we are inclined to give of names or, more generally, of singular terms, predicates, and quantifiers in the mother tongue [ought to] include those parts of the mother tongue we classify as mathematese... [3] I take it that we have only one such account: Tarski's. [Benacerraf 1973, my emphasis]

Benacerraf here adopts some fairly standard and seemingly innocuous views concerning natural languages and their interpretation. The picture is this: mathematical reasoning is carried out in some representational medium. The medium that math is 'done in' is, in essence, a precisified and disambiguated subset of our mother tongue.[90] We can agree to call this subset *'mathematese'*. Whatever theory of meaning we are inclined to offer for natural languages should extend to cover mathematese. In particular, whatever semantic treatment we are inclined to give noun phrases of the vernacular should be identical to the treatment that phrases which designate mathematicalia ('seventeen') receive. And since there is really only one sufficiently well articulated semantic theory that can do the job—the Tarski-inspired formal semantics discussed in Chapter 1—it is this sort of semantics that should be used so as to interpret mathematese.

In spite of its apparent plausibility, much of the story is wrongheaded. On reflection, and taking into account what we currently know, all three elements face significant challenges.

---

[89] Compare here a platonist philosopher's realism about the set of all English sentences and a modern syntactician's (equally robust) realism about the grammar that yields just the observed utterances, *modulo* production errors. See, for example, Katz [1998] and Chomsky [1986].

[90] I read [1] as claiming that mathematics is literally done in the vernacular in the sense that we psychologically process mathematical ideas by availing ourselves of our natural language. However [1] can perhaps also be read as claiming that we merely *state* mathematical judgements in NL while reasoning about them in some other, language-independent way. As a reading of the above passage, the latter interpretation seems dubious—how can the semantical apparatus of mathematics be part and parcel of mathematese if the latter is merely used to state results? Incidentally, an advocate of an language-independent access to mathematical content owes a concrete account. I offer mine in Chapter 4. (I'm grateful to Rob Stainton for drawing my attention to this ambiguity in Benacerraf's wording.)

To begin: the explicit presupposition on which [1] rests—i.e., that mathematical reasoning is 'done in' natural language—is inconsistent with results from the previous two chapters. Just a moment ago, we saw that even moderately sophisticated mathematical operations (factoring large numbers, locating primes, multiplication and division) can be accomplished by human subjects who lack any grasp of the syntax and semantics of natural language. If that's indeed correct then mathematical reasoning and linguistic competence dissociate. And so, it's very unlikely, on empirical grounds, that the representational medium employed by the thinking mathematician is her native tongue.

Indications of this were already implicit in Chapter 1. Mathematicians speak a variety of languages. Yet they apparently share a common mathematical reality and explore the same mathematical landscape. The natural language a mathematician happens to speak apparently has no impact on her adult mathematical skills or whether she is able to share her results with colleagues. Strictly, this is consistent with the existence of a *mathematese*: perhaps mathematical work is conducted in a shared subset of all vernaculars. Provided the subset employed lies at the intersection of all human natural languages (in some sense to be determined) we would not expect any special difficulties of translation to arise.[91] The trouble is that this hypothetical shared subset of the vernacular has further unusual properties. First, it cannot be rendered without loss in any of its supersets. Even moderately sophisticated mathematical results cannot be stated with full accuracy without resorting to purely mathematical concepts and mathematical notation. This can perhaps be explained away; poetry is exceedingly difficult, if not impossible, to explicate in prose as well. (The difference, of course, is that truth-conditional equivalents of lines of poetry are trivial to find; what is washed out by these is tone and poetic colour. By contrast, the truth-conditional equivalent of the claim, say, that every infinite binary tree contains an infinite path that makes no reference to trees, branching, posets, maximal elements *or other mathematical notions* is hard to fathom.) Furthermore, as we saw in our opening chapter, the syntax of mathematics (but not of poetry) and the concepts employed by it are indispensable to abductive inference in natural science. And this second property doesn't square at all well with *mathematese* being a subset of NL. Languages are part of our contingent, human endowment. Even if we allow that they share a common structure, it's hard to see how an overlapping subset can underwrite creative scientific abduction. This, together with the results from Chapter 2, should make us reject the presupposition on which [1] rests.

Even if the first of Benacerraf's claims is false, [2] can be read as directing us to an interesting issue—viz., the hypothesis that whatever semantic account one is inclined to give of declarative sentences in general should extend unmodified to statements that express mathematical judgements. Whether this is correct is a problem to which I will return in the next chapter. But before dealing with that topic, we face a logically antecedent issue.

[91] For a variant of this idea see Hauser et al. [2002].

Namely: what, in general, is the appropriate manner of developing a semantic theory for natural languages? In [3], Benacerraf identifies Tarski-Montague formal semantics as the only serious contender. Certainly, this remains the dominant approach—at least among philosophers. Earlier, we saw Saunders Mac Lane default to it as well. Nevertheless, I'd like now to argue that this is an issue well worth reconsidering. Formal semantics is no longer the sole contender. Moreover, FS faces some persistent and fundamental difficulties that call into question its explanatory power. I'd like now to turn to these with a view to showing that [3] is not correct.

Limitations of formal semantics

Scientific theories serve a variety of purposes. Ideally, we want a mature theory to classify, measure, locate causes, predict, and to explain. Here I will focus mainly two of these objectives: classification and explanation. There is mounting evidence that formal semantics succeeds as a taxonomic enterprise but that it does significantly less well as an explanation of the meaning of natural language expressions. If that's right then the door remains open for an alternative paradigm to step in and attempt to do better.

Let us say that a theory is *descriptively adequate* to the extent that it delivers an accurate, nuanced, perspicuous and (with luck) complete classification of a target domain. According to one time-honoured metaphor, good descriptions succeed in 'carving nature at the joints' and laying out the parts neatly. Ideally, we want a successful taxonomy to be internally coherent but also to fit well with adjacent domains (or at least not to clash). Developing a theory that manages to systematize a domain is a nontrivial achievement. Indeed, the history of science shows that classificatory adequacy is often only achieved after a number of false starts. Kuhn [1970] points out that in pre-paradigmantic sciences it's far from obvious which properties of the target domain researchers ought to attend to. For this reason, typically, early works include descriptions of trivial regularities liberally interspersed among phenomena that, in due course, come to figure in the laws and explanations of the mature enterprise. Before the situation has clarified itself, two or more rival classifications can compete for the affections of working investigators.

Among formal semantics' notable successes is the development of the first descriptively adequate classification of the meanings of natural language expressions. As we saw earlier, FS offers a finitely characterized, formally precise taxonomy of the meanings of an unbounded range of sentences, phrases and words. It captures the fact that meanings of complex expressions typically depend on the meanings of component parts and the order of their occurrence. And it sometimes also classifies sentences such that semantic relations such as synonymy, antonymy and logical entailment are highlighted. As we saw earlier, FS helps itself to a toolkit that enables it to meet this challenge. Here, again, is Montague:

> There is in my opinion no important theoretical difference between natural languages

> and the artificial languages of logicians; indeed, I consider it possible to comprehend the syntax and semantics of both kinds of language within a single, natural and mathematically precise theory. [Montague 1974c]

Viewing natural languages as essentially on par with artificial ones licences the use of analytic tools developed for the semantic interpretation of formal calculi. Central here is the notion of a model. A model $\mathfrak{A}$ of a language $\mathscr{L}$ is an ordered pair consisting of a domain D, and an assignment function which maps each name in $\mathscr{L}$ to an element in D. The model in the case of natural languages is just the domain of our natural discourse: all those entities which figure in our discussions. The assignment function is the reference relation that maps words to world and (as we saw earlier) recursively builds up interpretations of phrases and sentences. The end result is a flexible, systematic, and infinite hierarchy of semantic kinds—one that is capable of classifying a rich portion of any natural language.[92]

The sheer descriptive power of formal semantics is indisputable. Let me concede that FS is capable of offering a maximally nuanced nomenclature of semantic types such that any descriptively adequate theory of NL meaning can be restated in FS terms. What's less clear is whether FS constitutes an explanation of the phenomenon of linguistic meaning. Here we need to tread carefully. It's important not to stack the deck against FS by adopting a loaded conception of explanation. Just what ultimately counts as a ground-level explanation is a fraught issue. Outside of scientific contexts, we are typically quite tolerant concerning what we are willing to accept. The explanatory devices we employ depend on our goals and the level of precision sought: explanations of people's actions are typically intentional (in Dennett's [1998] sense). Sub-personal explanations and those that concern the inanimate are often functional ('because the kidneys filter blood') or causal ('because it struck the window'). Others still can be etiological ('because it comes from Persia'). Or they can be hybrid. Scientific explanations though tend to be significantly more regimented. Let me follow Hempel [1962] here: let us agree that to fully explain an *explanandum* is at least to show how it can validly be deduced from a set of covering laws. Let us agree also that everyday informal explanations are, in fact, elliptical versions of arguments that demonstrate how observed phenomena can be inferred from the natural laws. Nothing I want to say hinges on the peculiarities of Hempel's account. But accepting the deductive-nomological model of explanation prevents us from building in the requirement that an explanation necessarily specifies relationships of cause and effect. My hope is that we have here a characterization of explanation that everyone can work with for the moment.

At first glance, the word-world relations that formal semantics highlights look taylor-made to act as premises in deductive-nomological explanations—hence to act as explanations

[92]That portion of NL which involves tense, referentially opaque contexts, or other varieties of context-sensitivity require us to move beyond a purely referential semantics to one that includes *intensions*. I'm bypassing those issues here.

of meaning. Nonetheless, a number of linguists and philosophers—many of them working in a broadly cognitivist, Chomskyan tradition—have recently expressed doubts that this is so. In what follows, I draw on their work—especially the work of James McGilvray [1998], Paul Pietroski [2003a, 2003b], and Robert Stainton [2006a], as well as that of Noam Chomsky [2000]—to argue that, on closer scrutiny, FS really presupposes rather than furnishes an explanation of the meanings of natural language expressions. If this is the case then the conception of meaning upon which Benacerraf predicates the narrow version of his problem is flawed.[93]

Referents. As we just saw, the FS explanation of the nature of meanings depends ultimately on referential mappings between linguistic items and everyday objects. On closer scrutiny it turns out however that these word-world relations have some puzzling features which prevent them from appearing in the premises *bona fide* deductive-nomological explanations. Consider the following example, due to Chomsky [2000]:

(6) London is so unhappy, ugly, and polluted that it should be destroyed and rebuilt 100 miles away. [p.37]

Let's suppose for the sake of argument that this sentence is true. (If we want to insist that it's false, the same point can be made using an alternative example.) On the FS account, for this to be the case, at a bare minimum, there must be *an entity in the domain* that corresponds to the proper noun that begins the sentence. What sort of entity is it? Well, for a start it must be unhappy. It also needs to be ugly. And it must be polluted. Furthermore, this same entity needs to somehow be susceptible to being obliterated and recreated some distance away while retaining its identity. On reflection, all this is rather odd. What manner of natural object, after all, could simultaneously possess all of these traits? Chomsky comments:

> London is not a fiction, but considering it *as London*—that is, through the perspective of a city name, a particular type of linguistic expression—we accord it curious properties:. . . we allow that under some circumstances, it could be completely destroyed and rebuilt somewhere else, years or even millenia later, still being London, that same city. . . We can regard London with or without regard to its population: from one point of view, it is the same city if its people desert it; from another, we can say that London came to have a harsher feel to it through the Thatcher years, a comment on how people

[93] I've always thought it odd that we can allegedly successfully refer to entities outside our light cone. This peculiarity does not however constitute a full-blown objection to formal semantics unless one also accepts a causal theory of reference—something neither the formal semanticist nor a neutral arbiter have much motivation to do. Nevertheless, for the record, I do think that attempting to offer serious scientific explanations in terms of acausal relations is just short of invoking magic. Nothing in what follows is predicated on my prejudice. Some of my reasons were offered in Chapter 2.

> act and live. Referring to London, we can be talking about a location or area, people who sometimes live there, the air above it (but not too high), buildings, institutions, etc., *in various combinations.* [Chomsky 2000, added emphasis]

Try this thought experiment: Imagine we wished to explain to a nonhuman intelligence how some of our best scientific accounts represent the world. If our interlocutor were sufficiently clever and equipped with an appropriate sensory array and measuring instruments we should have no trouble conveying that (say) there is a property we call *momentum* and that it's a vector product of mass and velocity. Likewise, we could explain that we take the gravitational force exerted between two bodies to be proportional to the product of their masses divided by the square of the distance between them. In fact, most if not all of our physics, chemistry, biology, and other mature sciences could, plausibly (with effort) be made intelligible to our visitor. Now, what makes this possible is that in each case, the fields quantify over objective features of our shared reality or over complexes of such features. The point of Chomsky makes in the passage just quoted is that (barring a miraculous coincidence) the city of London would be a posit that would, in all likelihood, remain perpetually incomprehensible to our alien interlocutor. There is no single, coherent, natural entity existing independently of human concerns that counts as London. Let me be clear: The problem is not that 'London' is vague or ambiguous. Rather, it's just that the complex we pick out with this word is not a freestanding entity in the world but rather the product of our embedded, culture-dependent concerns and interpretations. In the end, our visitor would be forced patiently to learn by rote what human beings' semantic intuitions were regarding the correct extension of this term. (The air in the city? How high?)

I'll say more about the implications of this observation shortly. First, let's note that the issue is not restricted to proper nouns and their purported referents. Much the same can be said about common nouns. Here's an example:

(7) The thin, blue book, weighing four ounces, standing second from the left, was published in 1759 and caused such a scandal in Paris that it was publicly burned.

On its face, the sentence concerns *a single* entity: a perceptually available object. One could imagine it being uttered by a proud host showing off her private collection. Evidently, if the sentence is true, the object must be thin and blue. As well, if the sentence is true then the object in question was indeed published in 1759 on which occasion it caused a scandal. To explain this fact it's tempting to invoke the type-token distinction. Perhaps it wasn't the very token but rather its type which was first published in 1759 and caused so much ruckus? Surely that's closer to the truth. But then what was burned? Clearly not the type; certainly not if types are *abstract* objects. And not this very token either. Evidently the sentence is not about a single object at all but rather about a collection of instances of a type. Fine. This collection then is presumably what is blue and thin? And weighs four ounces? That's not right either… And so it seems that there is no single natural object that can plausibly bear the properties being attributed simultaneously by a reasonable and apparently true sentence. Our human ability to interpret such sentences without so much as noticing the hermeneutic eddies and currents we effortlessly navigate speaks volumes about our subtle interpretive capacities.

The problems multiply. Adjectives present parallel difficulties. Consider the following sentences:

(8) The apple was red.
(9) The traffic light turned red.
(10) Adam's house is red.
(11) Charlie's bedroom is red.

In Chapter 1, we saw that properties can be construed as unsaturated functions: on the standard account the relevant function maps to the true just in the case that the object it takes as an argument is, in fact, red. You and I understand this perfectly. But spare a thought once again for our alien intelligence. What *is* it to be red? Is it possible to offer objectively accessible, sufficient conditions for counting as red? Perhaps to be red is simply to be a surface that absorbs most wavelengths but reflects light between 635 and 740 *nm*? Not so. The redness of the light mentioned in (9) is not due to its subtractive but to its additive properties: rather than absorbing most wavelengths, a coloured light emits only electromagnetic radiation in a certain range. It might seem therefore that we need a disjunctive definition of redness in terms of a number of physical characteristics. Things are more complicated still.

> A ripe apple looks about the same under the whitish light of the full moon, the bluish light of the sky, or the yellow light of an incandescent light bulb. This is true even though the wavelength composition will differ significantly in all three cases. [Koch 2004, 137]

An account in terms of naked physical properties wavelength is clearly not enough. An account of redness that we offer our visitor will need to treat it not as a brute property but rather as a relation between an object, its surroundings, and our perceptual apparatus.[94] There are further difficulties with interpreting a colour predicate however. What counts as red is partly dependent on the kind of object being considered. Consider Adam's house, in

[94] Herman Weyl [1934] points out that "the sun is 150 000 times as bright as the full moon; consequently white paper under moonlight is darker than black velvet under sunlight; nevertheless, we see that paper as white, and not as black paper, both by moonlight and sunlight." The same point is brought home forcefully by the various colour illusions, including Adelson's checker-shadow illusion. See also the work of Kathleen Akins.

(10): it is decidedly *not* red if only its interior walls and its roof have have this colour. It *does* however count as red if the exterior walls have been painted (even if the roof is brown and the interior walls beige) provided some further caveats are respected. The red paint cannot constitute a pattern, such as stripes, checkers or polka-dots, for example. There are a variety of ways that a bedroom can be red. It might have red furniture (and neutral, eggshell walls) for example. What I take all this to show is that recognizing an object as red depends inextricably—and in surprising ways—on what human beings and human cultures are like. Once again: the point is not that there is no such thing as redness. Rather it's that being red is a property only recognizable from a certain perspective: ours. A semantic account that simply says that 'red' means red is of no use to an intelligence that does not already possess our prejudices and cognitive constraints. Rather than explaining how human beings understand their language, FS presupposes such understanding.

Words. Evidently, it's often hard to make sense of the right-hand side of the word-world relation, except *vis-à-vis* human cognitive faculties and social conventions. Many of the apparently freestanding props in 'the passing show' turn out, on closer scrutiny, to depend on us for their identity conditions.[95] The structure we are dealing with is therefore at least a three- rather than a two-place relation: viz., ⟨word, ⟨sensory input, perceptual and cognitive processing⟩⟩. The sensory input is recognized by us as a stable 'object' and a word is matched to that object. This is not the end of the story however. It turns out there are additional complications on the left-hand side of the relation as well. The individuation of words also relies essentially on human propensities and abilities.

The same word can be spoken, calligraphied, hand-signed, and telegraphed. In some cases, a word can be communicated in just the same way in English, Dutch or German.[96] So apparently, the identity of words transcends changes in medium and sometimes even in dialect. I propose that we bracket those complex issues for a moment however and restrict ourselves to the paradigmatic case: token words of a single dialect spoken aloud. I propose we inquire into what the identity of words thus construed consists.[97] Initially, it's tempting to look to the physical characteristics of the acoustic signal for an answer. Spectrograms provide a useful graphical illustration of that signal, a sort of 'voice print.' We can see by looking at the energy pattern however that the acoustic signal of human speech is continuously variable, and not, as one might perhaps expect, conveniently segmented at word boundaries. Indeed, one reason why spectrograms have not delivered the anticipated

[95]The exceptions, interestingly, are scientific posits, such as *nitrogen* and *energy*. These, Chomsky [2000] argues, are importantly different.

[96]The most obvious examples are certain geographic terms or names of monetary currencies. But many other borrowed words are homophonous as well.

[97]My discussion here draws heavily on Ingram [2007] and also on Jackendoff [2002].

quick strides in language recognition is that in order to identify lexical items in the signal, finer-grained elements within the energy patterns need to be matched with appropriate phonological representations—those phonological representations, in turn, can used to access appropriate lexical items. These are however nontrivial tasks. Matching acoustic signal to phonological target is complicated by a number of factors. Most obviously, there are individual differences between speakers due to differences in their vocal tracts. As well, the local speaking environment often changes how people use their voices and hence what the spectrogram registers. We often alter how we speak depending on social context—our sociolect can be more or less crisp depending on whether we are aiming to sound formal or casual. Finally, there are more technical reasons for variation. The pronunciation of individual sounds is often affected in subtle ways by what precedes and follows them. These coarticulation effects stymy any simple mapping between features of the physical signal and the phonological targets.

Let me say something about the phonological targets themselves. As I have already intimated, these representations stand in a one-to-many relationship with phonetic representations; they are, in a manner of speaking, abstract or (better) under-specified. Unlike the acoustic signal, which is continuous, phonological features are categorical. Moreover, they are hierarchically nested: at the top of the hierarchy, there are, phonological representations of entire words. The parsing of the phonetic signal may sometimes involve matching its components directly to such relatively corse-grained representations. Frequently however, the phonetic signal is matched to more fine-grained data structures that encode syllables as well as their various components. That this occurs is perhaps most easily seen in the case of nonsense words. Once the phonetic signal is matched to a phonological representation, the further problem of matching that representation to appropriate syntactic and semantic features needs to be solved. More on that in a moment. (See also Jackendoff [2002] and Ingram [2007].)

The upshot from the foregoing is that what common sense identifies as 'words' are not, in fact, freestanding entities (or even event tokens) discernible solely in terms of their intrinsic physical characteristics. The three-place relation I earlier proposed should, properly speaking, be expanded further. Its left hand side should read: ⟨⟨acoustic signal, phonological parsing⟩, lexical retrieval⟩. This yields a lexical representation complete with phonological, syntactic and perhaps some additional features (more on this shortly). The right-hand side of the relation remains this: ⟨sensory input, perceptual and cognitive processing⟩. Here, the retrieved lexical item is paired with the result of nonlinguistic perceptual and cognitive processing. The entire picture nicely explains our common-sense notion of a 'word.' But notice that words have here become an explanandum rather than the explanans. The individuation conditions on spoken words make essential reference to human recognition capacities. In this regard, the property of being a word turns out to be not unlike being red, being a book, or having the property of being the city of London.

Discussion. Here's what I take our discussion to show. A scientific theory faces a number of simultaneous challenges: among them, descriptive and explanatory adequacy. The explanation of the meanings of natural language expressions offered by traditional formal semantics relies crucially on mappings between freestanding linguistic items and the elements of a non-lingusitic domain—that is, on referential mappings between words and world. The evidence I have reviewed strongly suggests that this picture is too simple. If that's the case then FS meaning postulates ('ready' means ready, and so on) cannot play the same role in deductive nomological explanations as the objective regularities delivered by legitimate sciences. FS meaning postulates may, of course, still appear in DN explanations, but only as a sort of short-hand that itself stands in further need of elucidation by a more basic theory. To echo Paul Pietroski: rather than *offering* an explanation of the meanings of natural language expressions, formal semantics *presupposes* one.

Let me end the section by making clear why I do not think this argument is sufficient to demonstrate that the reference relation does not exist, full stop. The notion that words denote and that we use words to refer to things is central to our conception of ourselves as agents, language-users, and persons. It's perhaps no less central to our self-understanding than free-will or the propositional attitudes. What the arguments just presented suggest is that reference is no more (but no less) than part of our manifest image of ourselves; it's a part of our folk-linguistics. Reference is perfectly real. But it is only *visible* provided one adopts a sort of 'referential stance' (to use Dennett's useful phrase).[98]

In the final count, reference will not feature in a noncircular, scientific semantics. And so, for reasons wholly orthogonal to issues within the philosophy of mathematics, we must rethink the Benacerraf problem from the ground up. From here on in, I abandon the narrow version of the puzzle and pursue the wholly general issue of the meaning of mathematical statements and our cognitive access to their contents.

Linguistic internalism

Earlier, we agreed that an interpretation (or model) $\mathfrak{A}$ for a language $\mathscr{L}$ comprises a nonempty domain D and an assignment function $\mathscr{I}$ that maps the basic elements of $\mathscr{L}$ to elements in D. There is no reason to abandon this perfectly general, formal characterization. We do however need to critically reassess how we understand its three components: the

[98] The allusions to Dennett, Sellars [1962], and Chomsky [2000] are not gratuitious. For a particularly useful discussion of Dennett's own quasi-realist ascriptivism about propositional attitudes, see his [1998]. There one finds just this sort of quasi-realism that I think we should adopt toward the posits of folk-lingusitics, including reference.

language, domain, and assignment function in question. I now turn to that task.[99]

Language. The nature of natural languages came up briefly in Chapter 1, in the context of our discussion of mathematical realism. On that occasion, relatively little was said about the topic. I followed a prevalent philosophical tradition by assuming that human languages were something like rule-governed practices that communities engage in for the purpose of communication. At various moments, we also shifted emphasis from rules to sets and thus construed languages as sets of legal strings defined over an alphabet. Moreover, we followed Montague [1974a] in assuming that there was no theoretically important difference between interpreting formal and natural languages. Admittedly, we did make a perfunctory nod in the direction of linguistics by adopting a standard set of syntactic categories—IPs, NPs, VPs, and so on—though not much depended on that choice. Any finite fragment of a natural language can be viewed as having been generated by an unbounded variety of non-equivalent sets of instructions. In this context, getting NL syntax 'right' is simply a matter of offering some perspicuous description that facilitates the task of mapping sentences to truth conditions (words to objects, and so on).[100] The time has now come to reject this approach to the study of natural languages, their grammar, and their semantics root and branch. The alternative is to abandon a noncommittal (not to say instrumentalist) approach to NL syntax and instead take linguistic theory, along with its attendant philosophical commitments, much more seriously. In effect, it's time to extend our robust scientific realism to natural language.

Currently, the best available accounts of natural language fall within the Chomskyan paradigm. There are a variety of ways of characterizing that work. One way is to construe it as offering a solution to an important puzzle raised by the way that human children develop the ability to use and understand their native tongue. By the age of four, all children acquire a grasp of the local dialect (or dialects) provided some very minimal conditions are satisfied. All that's needed is that the child be neurotypical (within limits), that she be exposed to linguistic stimuli, and that she be permitted to sign or speak—in short, to participate. Children who meet these conditions achieve a stable grammatical competence very closely resembling that of their local community. Success does not depend on such factors as

[99] In this section, I mention *some* of the arguments in support of adopting the perspective on natural language that I favour. However, this is not the place for a full defense of the Chomskyan conception of language. The literature on the topic is vast and even a cursory review of the some of the controversies would take us off track. The unsympathetic reader is asked to treat what follows as an extended conditional: *if* one were to adopt a mentalist, nativist conception of language, such-and-such perspective on Benacerraf's problem and on mathematical realism would follow. For helpful discussions of Noam Chomsky and his intellectual legacy see Cook and Newson [2007], Hinzen [2006], McGilvray [2005], Smith [1999].

[100] Montague himself engages in rather careful syntactic analysis, though the categories he employs are now outmoded. Even so, Montague's [1974b] discussions of syntax as a pursuit are often dismissive and disparaging: "I fail to see any great interest in syntax except as preliminary to semantics."

the quality of care children receive (barring grotesque neglect), their general intelligence, attention span, memory, or any other cognitive or perceptual peculiarities [Yamada 1990]. This is odd for several reasons. Language acquisition occurs at an age when solving formally simpler problems—such as learning to play bridge or to do long division—is beyond a child's cognitive powers. By contrast, several decades of sustained collective labour by adult linguists (some of whom *can* play bridge) has not fully uncovered the principles governing even a single natural language. A second peculiar feature of the situation is its specificity; no other animal species has the capacity to learn the grammar of a human language in spite of the concerted efforts of primatologists and psychologists [Petitto 2005]. Evidently, the human child's rapid, accurate and apparently unique ability to grasp the grammar of her dialect calls for some sort of explanation.

Everyone agrees that our species' linguistic abilities are derived in some measure from a natural *endowment* and in some measure from exposure to our linguistic and perceptual *environment*. If there is a controversy here, it's over the relative contribution of each. The simplest explanation—the null hypothesis, if you like—emphasizes the role played by nurture and minimizes the contribution from our innate, biological constitution. On this sort of story, every human child comes into the world equipped with powerful, domain-general pattern-recognition capacities and the ability to mimic linguistic stimuli. Thus, a child learns her native language by committing to memory exemplars of adult linguistic constructions and subsequently reproducing *similar* constructions (perhaps replacing lexical item for lexical item). As it turns out, there is now considerable evidence that, in spite of its initial plausibility, the empiricist hypothesis is false.[101] Some of that evidence is negative; it comes from observing what children *never* do despite what empiricist models predict. Consider the following sentences:[102]

(12) The fact that Euclid had been superseded was not a surprise to some of the geometers.
(13) The fact that Euclid had been superseded was not a surprise to any of the geometers.

(14) The fact that Euclid had not been superseded was a surprise to some of the geometers.
(15*) The fact that Euclid had not been superseded was a surprise to any of the geometers.

The four sentences are structurally nearly identical. Nonetheless, every native English

[101]Empiricism fails here and elsewhere too. For evidence against empiricist accounts of arithmetic learning see Wynn [1992].

[102]The discussion here follows Pietroski and Crain [2002] and Pietroski and Crain [2005]. I am avoiding the hackneyed example of verb-fronting in English question-formation since an empiricist account of this phenomenon based on word frequency effects may arguably be possible. Pietroski and Crain's examples show that, even if this is so, poverty of the stimulus arguments remain compelling.

speaker recognizes (15*) as unacceptable. Evidently, an adult speaker relies on some rule G which allows her to generate the first three sentences and prevents her from generating the fourth, even though G is hard to make explicit. The challenge for a linguist is to spell out G and to explain how it comes to be known. In this case, the appropriate rule is this:

> G: Downward entailing linguistic environments license negative polarity items (NPIs) when a negative element c-commands the NPI.

Let me unpack that a little. A downward entailing linguistic environment is one which licenses inferences from a set to its subsets. If all geometers love circles then all Sicilian geometers love chalk circles. However, if only some geometers love circles then it's not necessarily the case that any Sicilian geometer loves a circle or, for that matter, that any geometer loves a chalk circle. 'All' therefore creates a downward entailing environment in both its argument positions; 'some' does not. And so, in the above example, (13) and (15*) count; (12) and (14) don't. Furthermore, negative polarity items are words like 'no', 'none', 'any', and 'ever.' These words can sometimes, though not always, appear legitimately in downward entailing environments. The challenge is to discern what features of sentences make the crucial difference between the licensed and the unlicensed appearance of an NPI. One logical possibility is that NPIs can appear in downward entailing environments so long as they are preceded by a 'not'. We can easily see from the above examples that proposal cannot be right; if it were, (15*) would be acceptable. In fact, the relationship that 'not' must bear to 'any' is somewhat more complex. To a good approximation, a phrase P c-commands another phrase R when R is among the daughter nodes of P's mother, and P and R are distinct [Ouhalla 1994]. In (13) the negation is part of the VP and c-commands the complement phrase in which 'any' appears. By contrast, in (15*) the negation appears in the NP and thus does not c-command elements of the VP. And this accounts for the acceptability of the former but not the latter sentence.

Evidently, the adult native English speaker somehow comes to be 'aware' of this rather baroque dependence. A child trying to learn the appropriate rule can go wrong in two ways: she can fail to generalize appropriately and, for example, mistakenly judge (13) to be ungrammatical. Conversely, she can overgeneralize by erroneously accepting (15*) as correct. Neither scenario is realized. Children do manage to grasp G. What's more, they do so without the benefit of correction or explicit instruction. This is remarkable since c-command is an obscure structural feature of NL syntax; it cannot be identified on the basis of surface features or lexical regularities. The empiricist therefore is left to explain how children arrive at G solely on the basis of the positive evidence, while not making overgeneralization errors.[103] No available empiricist model is able to explain this—or indeed for a host of

[103] Children do, of course, make mistakes when learning their native language. But such mistakes typically concern lexical items and not language universal aspects of grammar. More on errors in a moment.

similar examples of the acquisition of hidden linguistic structure. So long as this state of affairs persists, the empiricist null hypothesis is dead in the water. (While on the topic, let me note some recent positive evidence against empiricism as well. Pietroski and Crain [2002] discuss some of the grammatical errors that children *do* in fact make. It seems that while children do not erroneously overgeneralize from local grammatical constructions, they do sometimes produce constructions which would be legal in other languages. Thus, child learners of English produce constructions legal in Italian, for instance. Another challenge for the empiricist is to account for this regularity.)

Faced with evidence that external stimuli are too impoverished to account for the observed pattern of language acquisition, the Chomskyan reverses the explanatory priority of nature over nurture. If structural relationships crucial to our linguistic knowledge—such as c-command—cannot be abstracted from ambient linguistic data, they must in some sense form part of our innate linguistic endowment. On the Chomskyan conception, one important aspect of the linguist's job is to characterize the initial state of our linguistic knowledge (the *universal grammar* or UG) and to explain how UG interacts with the environment to yield the linguistic competence of the mature native speaker. We can be sure that UG is indeed universal since all human children are capable, in principle, of learning any human language: Welsh children brought up in Montréal learn Québécois; Czech children brought up in Barcelona learn Catalan. We can also be confident that the language faculty constitutes a discrete cognitive organ since a double-dissociation between our linguistic abilities and other mental functions can be demonstrated. Natural and induced aphasias provide us with examples of the selective impairment of the language faculty that can leave other abilities unimpaired [Berthier et al. 1990, Lecours and Joanette 1980]. Likewise, there exist instances of general cognitive deficit with spared linguistic abilities [Yamada 1990]. There is neurophysiological evidence as well. As we saw earlier, the language organ of some 90% of right-handed subjects is localized in Broca's and Wernicke's areas, so in the left hemisphere inferior frontal and superior temporal lobes Pulvermüller [2002]. Language, in the technical sense that interests the linguist and the cognitive scientist is thus an internal, individual, innate mental organ; it is *not* an abstract set of utterances or a collection of community practices and conventions.[104]

On one standard Chomskyan account, the initial state of the language organ comprises a set of parameters—or switches. One parameter may concern whether the subject of a declarative sentence can legally be dropped (as in Spanish but not French). Another parameter may concern whether the word-order in a sentence is relatively free (as in Greek but not English). Yet a third may determine whether a sentence can consist of *subject-verb-object* or *subject-object-verb* (Polish versus Japanese). On the assumption that parameters are binary

[104]This sheds some light on why such folk-linguistic posits as 'word' were earlier found to be explananda rather than explanans.

and that there are two dozen of them, there are over sixteen million (2^{24}) possible human languages—or, if one prefers to individuate differently, there is a single human language which admits a vast range of variation. The major task a child faces when acquiring her native tongue is to determine how the grammar of the local idiolect is parametrized. This process is not a matter of hypothesis formation and testing. Rather it's a matter of the setting (or prioritizing) of dozens of constraints.[105] This is not a trivial matter, but it *is* vastly easier than attempting to derive the correct regularities from scratch. (Note that this model accounts for why the grammatical mistakes children make are *correct* forms in other languages: a child making a grammatical error is trying out a parameter setting legal in another linguistic community but inappropriate in her local setting.)

Characterizing UG remains among the important long-term goals of linguistic research. In order to realize that goal, it's helpful to have sense of the possible final states of the language faculty; that is, to have a characterization of the language organ with parameters set and with lexicon fixed. This, of course, involves developing a theory of the linguistic competence of (some substantial subset of) mature adult speakers. Since parameter settings can be subtly different in each case, we focus on the individual. We also allow ourselves to abstract away from the inessential: we ignore all idiosyncrasies due to memory limitations, physical peculiarities, gender, age, and other extraneous, non-lingusitic factors. (Such idealization is not uncommon in the sciences; think here of frictionless planes and ideal gasses.) A successful syntactic theory (or grammar) is therefore a characterization of an internal, individual, cognitive capacity—a psychologically real process taking place in the mind/brain of the language user. In order to arrive at a grammar, all evidence is fair game: we draw on data from adult performance, children's language acquisition, native speaker 'intuitions' concerning grammaticality, cross-linguistic comparisons, studies of aphasia, data from second language acquisition, and neurophysiology (to name a few common sources). Notice how far this conception of linguistic research takes us from work on syntax in Montague's sense. While any number of grammars may be output-equivalent with the theory we are looking for, only one characterizes the actual cognitive process under study. We are not looking for *some* set of rules consistent with NL grammar; we're looking for *the* set of principles that characterize the structure of computations taking place in the mind/brain of the (slightly idealized) language user. Committing to a Chomskyan conception of language as a biological function (rather than formal structure) thus entails a shift toward a robust *scientific realism* in linguistics.[106]

[105] On an alternative account (which I prefer), language acquisition is a matter of determining the comparative importance of the various parameters [Kager 1999]. On this view every utterance violates *some* parameters; the unacceptable utterances however are sub-optimal with respect to which parameters they sacrifice. The precise version of the parameter account which we should adopt is not crucial for our purposes.

[106] See Smith [1999] for an extended discussion of Chomsky's work.

Domain. For reasons discussed earlier, the domain of interpretation for an I-language cannot consist of the set of perceptually salient properties and objects. Apparent referential mappings cannot underwrite semantic explanation; they are what needs explaining. And, in any case, we are capable of interpreting utterances about nonexistent, fictional, and even impossible entities—ones that *ex hypothesi* words cannot refer to.[107] There is a ready alternative: the relevant domain of interpretation for an I-language consists of information-bearing data-structures that exist outside the language organ—just beyond generative engine *per se*—yet that are nevertheless internal to the cognitive apparatus of the individual language user. This proposal can initially sound hopeless; it seems at first glance to entail a form of idealism. Nonetheless, it has recently been revived by several responsible philosophers and cognitive scientists—among them McGilvray [1998] and Jackendoff [2002].[108] In what follows, I follow their lead.

Let's begin by characterizing the 'meanings' of linguistic expressions functionally by attending to the roles they are called upon to play. Most obviously, meanings are what the words, phrases and sentences of natural language are used to express. Conversely, when all goes well, meanings are also what we derive by parsing others' utterances. Such interpretation is typically a matter of recalling relevant objects, places, situations, and events from long- and short-term memory. Sometimes, interpreting an utterance also requires that we take action so as to fulfill a request, heed a warning, or shift our perceptual focus. If we accept an internalist conception of language then whatever meanings ultimately prove to be they must act as a conduit between elements of the I-language and a host of other cognitive and perceptual subsystems—among these, the perceptual modalities, object recognition, motor capacities, episodic memory, and semantic memory. Evidently, language is not strictly necessary for some of the information integration at issue here. We are, after all, capable of listening to the same concerto that we see being performed. We are able to both see and smell a fresh loaf of bread while remembering the last occasion this happened. And we can decide to extend our hand so as to move a saltshaker that we imagine may fall off a table. These examples point to a language-independent ability to *associate* visual, auditory and olfactory stimuli with long-term episodic memory and with visual imagination [Hume 1748]. The linguistic interface however does make an important contribution. Elizabeth Spelke [2002] has argued, on the basis of several decades of experimental work, that the remarkable intellectual capacities of human beings can precisely be attributed to our flexible integration of information across a variety of domains, rather than to additional, domain-specific

[107]See Scott [2003] for evidence that the human conceptual apparatus treats real and imaginary objects in just the same way.

[108]McGilvray [1998] develops his ideas concerning semantic internalism within a minimalist extension of the principles and parameters framework; Culicover and Jackendoff [2005] pursue a different conception of grammar. The controversy over minimalism is not crucial for us here, so while I note it for the sake of accuracy, I will largely pass over it in silence.

cognitive powers. She suggests that these integrative capacities—capacities that are crucial for higher cognitive function, including the understanding of the logical operators—are tied to our linguistic abilities and come online as our language organ matures.

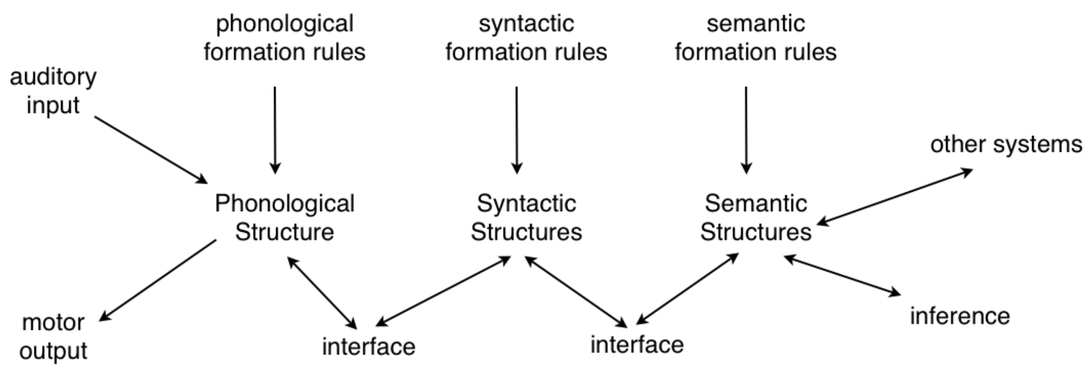


Figure 3: Jackendoff's [2007] parallel cognitive architecture.

The integration of information via linguistic meanings has a clear advantage over straightforward association. Among the striking features of I-language is its unboundedness; we can produce and parse an infinite number of distinct syntactic structures. And, of course, we can assign meanings to these structures. (It could have been otherwise. We might have had a generative grammar but only a finite list of interpretations we could assign to syntactic constructions.) Our capacity to construct meanings is therefore apparently *productive*. Moreover, much like syntactic constructions, meanings are *systematic*; someone who can understand that the cat is on the mat can also understand that the mat is on the cat. Our best—indeed, our only—explanation of productivity and systematicity is that we are dealing with a *generative* engine whose complex elements are composed recursively from a finite base of primitives by means of a finite set of formation rules [Fodor 1987]. A computational system of this type is significantly more flexible than an associationist engine.

Ray Jackendoff's [1992b] *conceptualist semantics* (hereafter CS) enshrines these insights. Jackendoff—whom we encountered in Chapter 1—rejects the traditional tacit assumption that NL syntax constitutes the sole generative module of the language faculty. In its place, he substitutes a parallel architecture comprising separate generative phonological, syntactic and semantic engines. This move not only has the advantage of explicitly recognizing the generative nature of our conceptual apparatus. It also squares well with the current state of research in linguistics. Phonologists (recall) work with a proprietary set of nested data structures that have their own rules of formation, composition and manipulation. Phonological representations are independent of the syntactic types used by grammarians, though they too are structured and so have their own local 'syntax' [Ingram 2007]. Many semanticists have also been relying on representations which form and combine independently of NL syntax [Fauconnier 1994, Talmy 2001]. Each of these is understood as working with its own data-types, having its own rules of formation and processing, and each is construed as interfacing with the other two. The result is an ecumenical model in which the phonologists', syntacticians' and semanticists' local theories can find their place (see Fig. 3). Each module possesses its own, dedicated rules for representation-formation and its own means of computing over those. The modules operate autonomously and in parallel but are related by a series of interfaces. Phonological processing of an input activates syntactic parsing. And this, in turn, occasions an attempt at semantic processing. If the processing does not terminate in error, an interpretation of the target phrase is constructed. Furthermore, since semantics (unlike syntax or phonology) is directly legible to mental operations outside of the language faculty, semantic interpretations are capable of triggering motor responses, episodic memories, spatial judgements, and a variety of other nonlinguistic processes. In effect, semantics becomes the language faculty's window on the rest of the mind/brain.

Conceptual structures play an important integrative role. It's advisable however not to build in more than we strictly need; meanings (construed now as cognitive interfaces) need *not* be representations in the philosophers' sense. When philosophers speak of *representations*, they typically mean symbols that re-present some worldly property, entity or event to a sentient subject. Thus understood, representations are essentially intentional; they point beyond themselves to an intentional object, either real or imagined. Indeed, representations are individuated (in part, at least) precisely with respect to their intentional properties.[109] The data-structures that act as interfaces between the I-language and other cognitive faculties should not be understood as representations in this sense. Syntactic structures emphatically *do not* refer to semantic structures. In fact, the reference relation itself drops out as a term of analysis. CS structures are not individuated by their intentional objects. Rather, they are individuated by their local, syntactic properties—viz., the cognitive, motor, and perceptual systems they interface with, and the local computations (inferences) the permit [McGilvray 1998].

Abandoning intentionality and the traditional conception of representation may seem unwise. But, in fact, there are some respectable precedents in cognitive science for just this move. There are even existence proofs: Rodney Brooks' [1991] robots are capable of complex interactions with their environment—even collecting pop-cans left around the lab—without the benefit of inner representations in the philosophers' sense. The activation vectors at

[109] The elements of Fodor's [1975] language of thought are representations in just this sense.

the higher levels of the robots' hierarchical subsumption architecture demonstrably do not correspond to features in the world at large.[110] The proposal here concerning interpretive interfaces that link the language organ to other cognitive areas is that they play a similar role. Rather than standing in need of further interpretation, cognitive interfaces *just are* the interpretations of linguistic terms. They perform all of the functions that we expect the meanings of our words and phrases to perform [Jackendoff 2002 p.306]. The inner information bearing states that syntactic elements map to are not representations in the sense that they do not 'stand for' anything. They are what McGilvray [1998] calls *intrinsic contents*. (Of course, it may turn out that, quite apart from their intrinsic properties, some of the interfaces do indeed additionally correspond to entities, properties, and events outside the head. Plausibly the meanings of the terms of completed sciences would have this feature. Any such correspondences would however be a fortuitous feature of the interfaces in question.)

Let me note another peculiarity of the CS account. The parallel architecture does not include a lexicon in addition to its three sub-modules. What we pretheoretically think of as *words* are really the result of the co-ordination of phonological, syntactic, and semantic computations. Under the hood however, there are only regularities in the way that the three modules interact.[111]

Assignment. Let's now turn to specifics. Ultimately, we need to exhibit a semantic interpretation of a sentence that expresses a mathematical judgement, and we can't do that without a clear sense of how, concretely, a conceptualist semantic analysis proceeds. We thus need to know more about the details of the assignment function that attempts to capture the (psychologically real) interface between I-language and conceptual structure. There are two major changes in this respect from our earlier FS account: CS relies on a vastly expanded repertoire of semantic types; and, moreover, nearly every syntactic category (not just NPs) are understood to be capable of being assigned to an element at conceptual structure.

The CS domain consists—not of freestanding things-in-themselves—but of all manner of stable posits that can arrest human attention. It's possible to categorize such 'things' and thereby to arrive at a catalogue of semantic types. This task is facilitated by adopting a two-pronged approach. A rough and ready catalogue can be arrived at quickly by reviewing the sorts of 'things' we can point to with our fingers. A roughly matching (though slightly

[110] Trained connectionist networks similarly do not manipulate explicit *representations* yet manage to accomplish complex classificatory tasks. See Churchland and Sejnowski [1992].

[111] This lets Jackendoff elegantly explain situations where the coordination comes apart slightly or fully. 'Pushing up daisies' and similar idioms have a normal syntax and phonology but interface as an atomic unit with a single semantic item, in this case: [DIE]. Likewise, nonsense phrases like "*twas brillig and the slithy toves did gyre and gimble in the wabe*" have a phonology and a syntax of sorts but fail to interface with any semantic structures.

more constrained) list can be obtained by considering the possible answers to questions such as these:

(16) What is *that*?
(17) What's *there*?
(18) What's *thataway*?
(19) *How*?

On reflection, we can see that the possible objects of attention constitute an ontologically heterogenous menagerie: entities, people, animals, properties, events, locations, amounts, times, paths, compass directions, and procedures can all become subjects of scrutiny. A desire for parsimony might incline us to attempt to reduce this list to a more basic list or even to a single, fundamental root—perhaps entities and events, or even entities alone would suffice. Jackendoff [1983] argues that this desire for simplicity is misguided. It's implausible on its face that events like signing a contract, or complex trajectories, or procedures like fixing a timepiece can be explicated without loss in terms of entities and their relations—except perhaps by means of an open ended and wildly disjunctive account. Indeed, this was the point of our earlier discussion about the city of London, colours, and books. It's not clear that adding events to our fundamental list helps.[112] Moreover, there is another reason not to engage in premature reduction. Ontological parsimony is not a legitimate constraint at this stage of research. Biological systems such as the human brain are, in some respects, inefficient organs where kluges and redundancies arise.[113] There is no *a priori* reason to expect that the interface between I-language and deeper, nonlinguistic faculties will answer to an ideal of computational efficiency. It's surely an empirical question which categories the human conceptual apparatus evolved to register independently. It's a further empirical question to what extent the cuts the brain evolved to make correspond to natural kinds. Rather than forcing a premature reduction, it's best for now to work with a multiplicity of types while awaiting guidance from ongoing work in cognitive psychology and neuroscience.[114]

Let's continue with our attempt at classification. Notice that many subsentential phrases are incapable of standing alone either as answers to questions or as freestanding assertions. Subordonate clauses such as "that Jack built" are an example. By contrast, phrases designating each of the conceptual types named earlier *can* quite sensibly be uttered alone. Here are a few examples along with their corresponding types:

[112] This remains an open question.

[113] Think here of our visual system, with its multiple and redundant pathways. There is no a priori reason that a projection to the superior colliculus should exist at all. Nevermind the what/where path divergence.

[114] For an excellent anthology of readings concerning conceptual structure, see Margolis and Laurence [1999].

(20)	Alisa.	⟨PERSON⟩
(21)	A pen.	⟨THING⟩
(22)	A misunderstanding.	⟨EVENT⟩
(23)	Two liters.	⟨AMOUNT⟩
(24)	Royal blue.	⟨PROPERTY⟩
(25)	Along the fence.	⟨PATH⟩
(26)	Politely.	⟨MANNER⟩
(27)	Next to the fridge.	⟨LOCATION⟩

Jackendoff [1983] suggests that fragments like (20)-(27) strike us as sensible standalone items because we can readily imagine cases where they would help fill a gap in conceptual structure. In each instance, we can easily construct a question, to which the fragment constitutes a full reply:

(28)	Whom did you see?	⟨PERSON⟩
(29)	What did she find?	⟨THING⟩
(30)	What transpired?	⟨EVENT⟩
(31)	How much?	⟨AMOUNT⟩
(32)	What colour?	⟨PROPERTY⟩
(33)	Which way did she come?	⟨PATH⟩
(34)	How did she say it?	⟨MANNER⟩
(35)	Where are my keys?	⟨LOCATION⟩

Notice two things here. First, each of these questions can be answered by pointing to the relevant entity, trajectory, location, or what have you, or by engaging in some creative mimicry. This suggests that in each case, we are dealing with a structure that can be encoded both linguistically and non-lingusitically. Note also the degree of specificity in the match between questions and answers. Attempted cross-category answers are nearly always nonsensical. "A pen" is not a readily interpretable answer to "Where are my keys?" for example. Each question demands a particular conceptual type as a possible answer. This once again suggests that many of the semantic types are mutually irreducible.

Let us accept the above list of semantic types as primitive. By doing so, we are effectively moving away from a semantics with a simple, unstructured domain. We are adopting instead a domain comprising a multiplicity of types. Although Jackendoff does not make this explicit, this is tantamount to using a many-sorted (or typed) logic to interpret I-language [Bell et al. 2001]. If we take semantic types to be mutually irreducible on grounds of biological realism, then the natural model structure for a natural language $\mathscr{L}$ comprises a sequence of domains $\langle D_0 ... D_n, \rangle$, where each domain corresponds to a major semantic category (path, event, etc.). The interpretation function $\mathscr{I}$ maps syntactic items into that ordered set of domains. How it does so brings up our second important difference between FS and CS.

Like FS, CS relies on function composition to construct a semantic interpretation. Relying on a multiplicity of irreducible semantic types changes what such an interpretations can look like. Consider, for instance, how NPs need to be treated on our new account. Earlier, we assigned singular terms occurring at the heads of NPs to entities of type ⟨e⟩. Now that our domain of interpretation is populated with a plethora of types, this picture evolves. As examples (20)-(24) show, many NP heads designate categories other than ⟨entity⟩. They are free to map to ⟨persons⟩, ⟨events⟩, ⟨amounts⟩, or nearly any other semantic type. There are parallel changes in how the heads of other syntactic types are treated as well. On the FS analysis, the heads of lexical categories other than NP are nonreferential; they map to unsaturated functions. CS by contrast takes *all* heads of phrases, regardless of syntactic type, to perform an equivalent semantic task by mapping directly to conceptual representations. Examples (25)-(27) illustrate this for PPs and APs. Here again, there are no straightforward regularities concerning the type of semantic structure a syntactic type can pick out. Finally, the maximal projections of phrases themselves play an important role. We will see in a moment that they form the exoskeleton structure within which an interpretation proceeds; each maximal projection is mapped to a distinct representation in conceptual structure.

Let me end this brief discussion of CS by mentioning one important similarity between it and FS. It concerns the limits of semantics. Neither theory wants to risk becoming a sort of catch-all theory of everything. In both cases, a great deal of information is understood not to be encoded in semantics. The idea is familiar from FS:

> We should not expect a semantic theory to furnish an account of how any two expressions belonging to the same syntactic category differ in meaning. 'Walk' and 'run', for instance, and 'unicorn' and 'zebra' certainly do differ in meaning... [but discerning how] demands considerable knowledge of the world. [Thomason 1974, p.48]

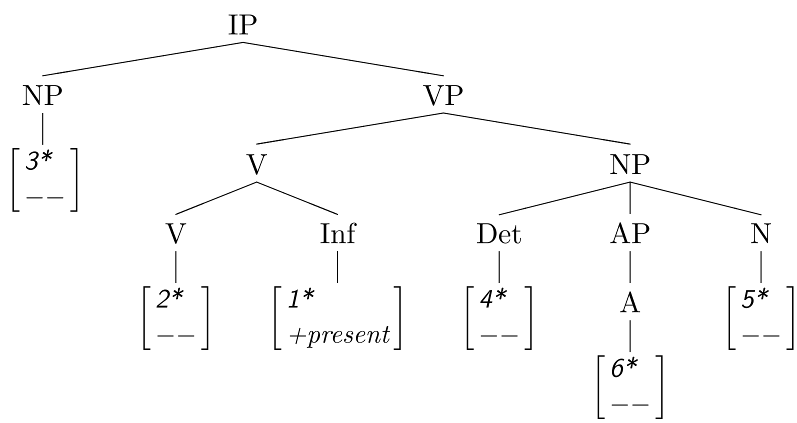
CS is no different in this regard. Conceptual structures are neutral with respect to a great deal of encyclopedic information concerning the ⟨entities⟩, ⟨events⟩, and other types they encode. In general, CS is an interface, not an encyclopedia. Whatever information is processed elsewhere in the brain need not be replicated at CS. Semantic structures therefore are not thought to capture the differences between chairs and seats, or between unicorns and zebras. *That* is left to long-term memory and to perceptual faculties. The point is worth making since it helps address an issue that was raised briefly earlier and set aside. Recall that Benacerraf hypothesized that the semantic interpretation of sentences that express mathematical content proceeds no differently from the interpretation of sentences that express any other factual states of affairs. In one sense, on the CS view, this is correct; both types of sentences are treated exactly on a par at the CS interface. Whatever differences there may be lie not at conceptual structure. They lie in what conceptual structures invoked by mathematical statements subsequently invoke further downstream in the cognitive architecture. (I will return to this topic shortly.)

Reanalysis. We can now put our revised semantic theory into practise. Our original goal (one half of it) was to offer a plausible interpretation of the *meaning* of sentences that express mathematical knowledge. We are now able to do this. Let's return to the sentence we have been working with throughout our study.

(36) Seventeen is a prime number.

In order to interpret it, we will first need to extract the sentence's syntactic structure. Here is one hypothesis concerning that structure:[115]

(37)



Several aspects of this tree require commentary. Most striking perhaps is the fact that the heads of phrases do not terminate in words. The reason is probably obvious given the foregoing discussion: 'words' are not, properly speaking, syntactic posits. They are folk-linguistic terms. The familiar appearance that there exist 'words' results from the co-ordinated interaction of the phonological, syntactic, and semantic engines. The heads of the phrases in this tree therefore contain (what are, in effect) pointers to corresponding structures in phonology and in semantics. I have labelled these pointers with starred numbers for the sake of clarity. The order of the pointers corresponds to the structural order of the constituents. The VP is taken to govern the sentence as a whole. The maximal projection of the first NP acts as the IP's specifier. The second NP plays the role of a complement to the verb

[115]The analysis is in line with Culicover and Jackendoff [2005]. Whether it is correct is, in the long run, an empirical question. I am using it here to make the fit between syntactic structure and conceptual semantics as perspicuous as possible.

phrase. It is taken here to project three elements which facilitate the explanation of the syntax-semantics interface (see below).

We saw a moment ago that, as a general rule, we can expect every major phrasal constituent to map to a major semantic category. We can expect the leaf nodes to map to functions (though these may be zero-place functions). Since the syntactic structure we're working with contains six phrasal constituents, we can expect the corresponding semantic structure to contain the same number of fundamental components. Here is that semantic structure:

(38)
$$\left[PRESENT \left[BE_{\text{ident}}\left(\begin{bmatrix} TOKEN : Seventeen \end{bmatrix}_{\text{Thing 3}}, \begin{bmatrix} TYPE : number \\ \begin{bmatrix} prime \end{bmatrix}_{\text{Property 6}} \end{bmatrix}_{\text{Indef4 Thing5}} \right) \right]_{\text{State 2}} \right]_{\text{Situation 1}}$$

The sentence as a whole expresses a situation which obtains in the present tense. This situation is represented in conceptual structure as involving the existence of a state between a token item and a type. The token is a ⟨thing⟩: the number seventeen. The type is complex: it involves being a number, also a species of ⟨thing⟩. But it is further delimited by the ⟨property⟩ of being prime. So the sentence expresses that the token entity named by 'seventeen' is found among the entities designated by 'number'; more precisely, it is among those entities of that type which are also prime. None of this is particularly surprising, of course. The usefulness of this analysis, again, hinges on the correspondence of the semantic types to biologically real structures and processes. Among cognitive scientists' research goals is to characterize these semantic constituents in the terms of computational neuroscience.[116]

The aim of the section was to explain how the meanings of sentences that express mathematical content can be understood in CS terms. We have done that. The new analysis is a mixed blessing: CS offers a fresh vantage from which to survey the general Benacerraf puzzle. You may recall however that some problems with Jackendoff's account of mathematical concepts arose already in Chapter 1. I now return to a discussion of those problems.

[116] For an early stick-and-ball textbook treatment, see Chapter 10 of O'Reilly and Munakata [2000]. For a very interesting attempt to offer a computational model of semantic structures similar to the ones we have been discussing, see Regier [1996]. For a recent discussion, see Jackendoff [2007].

Realism redefined

Tarski [2001] admonishes us that semantics is not a means of resolving substantive metaphysical disputes nor is it a way of showing that everyone except the semanticist and her friends are speaking nonsense. It's important to heed that advice. The best available accounts of natural language (syntax and semantics) fall within the Chomskyan paradigm. Adopting an internalist, mentalist characterization of language vitiates our earlier argument for mathematical realism. In a moment I will spell out why. What I want to stress however is that the overall metaphysical situation remains unaltered. Adopting a new semantic theory neither enhances nor diminishes the fundamental plausibility of ontological realism about mathematics. It simply calls for revisions.

Let's again turn to specifics. Earlier I argued that a minimal ontological realism about mathematics comprises two theses:

($\mathcal{R}$.i) Some mathematical entities exist; and

($\mathcal{R}$.ii) Their existence is independent of human minds, cultures, languages, and conventions.

You will recall that the argument for ($\mathcal{R}$.i) ran as follows:

1. Mathematical statements are true, false or lack content (hence are neither true nor false).
2. The abductive indispensability of mathematical statements entails that they are not contentless.
3. The deductive utility of mathematical statements shows that they are not false.
4. Some mathematical statements are *true*. [1,2,3]
5. Assuming traditional formal semantics, the existence of mathematical *truths* entails the existence of mathematical *entities*.
6. Traditional formal semantics is correct.

$\therefore$ ($\mathcal{R}$.i). Some mathematical entities exist. [4,5,6]

We now have grounds to reject (6); yet without it the argument does not go through.[117] It seems therefore that if we accept CS, we are left with no argument for NORM. The situation is further complicated by our independent arguments in favour of ($\mathcal{R}$.ii), NORM's second plank. A number of options suggest themselves. Most obviously we could reject the Chomskyan conception of language or conceptualist semantics. Evidently, the wholesale rejection

[117] Notice that this is a general result. Any philosophy that argues for the existence of mathematical entities on the basis of formal semantics thereby comes to owe a reply to the Chomskyan arguments presented earlier in this chapter. This extends to many platonist proposals, including Shapiro [1997].

of a scientific theory carries a heavy cost. We immediately come to owe an explanation of the phenomena that the theory successfully treats as well as an explanation of its apparent success. More modestly, we could try to derive ($\mathcal{R}$.i) from CS. This too, I think, is hopeless. CS does not posit the same language-world relations FS employs; tacking these on artificially would be *ad hoc*. Worse, it would be incoherent. A third option is to modify NORM itself. In other words, we can try to make sense of mathematical realism, mathematical facts, and mathematical objectivity without an appeal to mathematical *entities* per se. This seems like a reasonable project; the more so because there are precedents in the literature.[118] In any case, it would be peculiar were it to turn out not to be possible to articulate a form of ontological realism (about ion channels and electric charges, say) that was consistent with CS. Conceptualist semantics is, after all, no more than a hypothesis about how the human mind/brain initially interprets the grammatical structures made available by generative grammar; it's thus an internalist, computational theory of a particular biological process. Such a theory may arguably have some bearing on how human beings *construe* their surroundings and reason about them. But it would be bizarre indeed if a rarefied biological theory concerning the information-processing capacities of a particular organism were somehow to entail the non-existence of unrelated entities in the organism's surroundings.[119]

The key to articulating what ontological realism looks like under CS is to focus on the difference between more and less constrained cognitive processes. No mental process, including free-association and mental random-number generation, is *wholly* unconstrained. It's well known that human beings are particularly poor at generating truly random responses. The constraints operative in such situations are not particularly interesting however insofar as they do not reflect anything about a mind-independent domain. A more interesting case of constraints on cognitive processing arises in connection with sensory perception where, evidently, what we perceive is tightly—albeit imperfectly—constrained by the local environment.[120] Of course, not all constraints come from without. We cannot help but to recognize certain sentences in our native tongue as ungrammatical. In this case, we are constrained by our idiolect's parameter settings. Our judgement, as in the case of perceptual experience, is evidently not foolproof; garden path sentences can make us *think* that a constraint has been violated when it has not. This is no more troubling to

[118] Recall that in Chapter 1 I left the door open to alternative ways of construing realism. Putnam [2004] and Hellman [1989] offer a precedent for ontological realism without entities.

[119] This argument does raise certain interesting questions about the nature of our science-forming faculty. Chomsky 2000 suggests that this faculty constitutes an independent mental organ. Given the contrasting fortunes of natural language and mathematics in furnishing a useful vocabulary for natural science, the idea seems plausible. I will say more about this in Chapter 4.

[120] I leave a discussion of visual perception for another time. But see Biederman [1995] and Churchland et al. [1994] for two interesting perspectives.

the linguist however than visual illusions are to the vision-scientist. The correct syntactic theory accounts both for our intuitions of grammaticality and for why we find garden-path sentences so difficult to parse. Indeed, in each of the above cases, the fact that we can talk about making *errors* means also that we can talk about getting things right. And to claim that some representation is correct is (at the very least) to maintain that it does not violate any of the relevant, local, operative constraints. Evidently, 'correctness' is a very general notion; it applies to statements, procedures, choices, and so on. But when we limit our purview to contentful, affirmative judgements, to judge correctly is to judge *truly*. Bearing this in mind, here is our revised argument for realism:

1. Mathematical statements are true, false or lack content (hence are neither true nor false).
2. The abductive indispensability of mathematical statements entails that they are not contentless.
3. The deductive utility of mathematical statements shows that they are not false.
4. Some mathematical statements are *true*. [1,2,3]
5. Assuming conceptual semantics the existence of mathematical *truths* entails the existence of *constraints* on mathematical judgements.
6. Conceptual semantics is correct.

$\therefore$ ($\mathcal{R}$.i$'$). Constraints on (some) mathematical judgements apply. [4,5,6]

An ontological realist about mathematics can accept the above argument. She can add, moreover, that the constraints which delimit mathematical research are not due to the peculiarities of human physiology or psychology. They run deeper than any contingent features of our constitution. For a judgement to be universal, necessary and objective, it suffices for the constraints on that judgement to be mind-independent, inescapable, and applicable under all circumstances. We can now redefine NORM as follows:

($\mathcal{R}$.i) Constraints on some mathematical judgements apply; and

($\mathcal{R}$.ii) Such constraints' existence is independent of human minds, cultures, languages, and conventions.

The epistemic issue henceforth shifts away from how we 'make contact' with the truth-makers of mathematical expressions; instead we need to ask how the semantic structures employed in the course of mathematical reasoning are forced to follow a rigidly constrained course. What, in effect, makes the geography of peaks, valleys and hidden trails in Allain Connes' mathematical landscape so implacably resistant to our ambitions, desires, whims, and our wishful thinking? This will be our next topic.

Conclusion

It may be useful to recap and review. The problem we set ourselves was both to account for the semantics of expressions that purport to make factual mathematical statements and to offer an account of our knowledge of their contents. In Chapter 1 I argued that we have compelling reasons to opt for a realist approach to this problem—viz., an approach that takes the existence of (some) mathematical facts seriously. I argued in Chapter 2 however that our ontological realism about mathematics should not be a form of platonism; we ought not accept, in other words, that there exist acausal, abstract objects. That much is background.

I opened this chapter by pointing out that our initial argument for ontological realism (and indeed our initial statement of the problem under discussion) enshrined certain assumptions concerning the nature of human languages. In particular, we identified natural languages with publicly accessible conventions modelled by sets of well-formed formulae. We also took semantics to be a matter of mapping between linguistic items and freestanding entities out in the world. The thrust of this chapter was to argue that these assumptions are wrongheaded. Henceforth, we need to predicate our discussion on an entirely different conception of natural languages and their meanings. The place to look, I suggested, is modern linguistics. Chomskyan generative grammar offers us a powerful, realist account of the syntax of natural languages. Ray Jackendoff's conceptualist semantics supplements this with an attractive theory of meaning and an ecumenical parallel cognitive architecture. My claim here is in these theories we find half of the solution to the puzzle with which we began—i.e., we find a theory of meaning for linguistic expressions, including ones that express mathematical content.

The price we pay for this shift in semantic perspective however is that we must abandon our original argument for NORM; the existence of semantic representations of mathematical objects does not guarantee the existence of the corresponding entities *per se*. At the close of the chapter, I argued that we must nonetheless hang on to ontological realism. That's because our original argument reasserts itself in terms of *constraints* on mathematical judgements. In effect, we shift from an object-based ("extensional") realism about mathematics to a rule-based ("intensional") version, much as the linguist has shifted from construing languages as freestanding, abstract sets to viewing them as biologically instantiated functions. The crucial difference is that mathematics goes deeper. A minimal ontological realist about mathematics is one who recognizes the existence of mind-independent, objective, and necessary standards in the field—standards not dependent on human physiology or psychology. The remaining challenge is to specify the source of those standards in a way that allows for our mathematical knowledge.

4 The Epistemic Problem

Everything should lead back to...
objects in the realm of space and time,
and to lawlike relations that obtain
for these objects.
— Albert Einstein

Having now offered a hypothesis concerning that part of the *meaning* of sentences that express mathematical judgements which can be gleaned from the semantics of natural languages, it remains to account for our knowledge of their contents. Toward the end of the previous chapter, I suggested that mathematical content can be construed as a hidden ur-structure of nonnegotiable constraints operative on mathematical judgements. In this chapter, I'd like to offer the beginnings of an account of how those constraints arise. I will start by arguing that the relevant constraints cannot be accounted for by existing psychological or neurophysiological theories; psychologism, in other words, is false. We have the seeds of a workable realist alternative however in the writings of Kurt Gödel. My aim here will be to present a new interpretation of Gödel's philosophy and argue that, once charitably read and correctly understood, his work offers us an attractive explanation of how mathematical knowledge is possible and (equally importantly) also suggests a methodology by means of which we can hope to supplement and extend ongoing work in cognitive science.

Psychologism

Let us suppose on the basis of the discussion in the last chapter that lexical items which encode mathematical content are partially constituted by their semantic interface. (Whether *this* specific theory or some Minimalist alternative ultimately proves correct is not essential.) The semantic interface itself contains no special mathematical information; it merely mediates between the language faculty and other mental organs. The cognitive psychology and cognitive neuroscience literature are therefore the obvious place to look for an account of the mathematical content that semantic structures encode. Those fields are now converging on several models that promise eventually to explain the information processing involved in our mathematical (and particularly arithmetic) reasoning. Neuroimaging and neurophysiological investigations are furthermore starting to correlate psychological models with the cortical regions where the arithmetic processing is thought to take place. I'd like to discuss some of that work. In Chapter 2, I argued that philosophical accounts of mathematical knowledge ignore empirical results at their peril. Now it's time to see the other side of that coin. Scientific hypotheses concerning our knowledge of mathematical facts which sideline contributions by mathematicians (and also philosophers of mathematics) can also be expected to come up short. It will be instructive to see why that is.[121]

Stanislas Dehaene and his collaborators [2003] suggest that human arithmetic calculation involves three distinct *kinds* of processing. At its core, it involves the processing of purely *quantitative* information. In certain circumstances, additional *linguistic* processing is also involved. Finally, some activities also require the use of an additional *visual* code. Each of these activities can be correlated with the actions of specialized circuits located in the parietal lobe. As one might expect, given the vast complexity of the brain, the story is not entirely straightforward: the fit between the cognitive psychology and neuroscience is not wholly perfect; and, moreover, several regions outside the parietal seem also to participate in mathematical reasoning.[122] Nonetheless, the models and results are robust enough that it's possible to offer a preliminary outline of the 'mathematics faculty.'

Bilateral intraparietal. Dehaene *et al.* [2003] suggest that the core quantity system is localized bilaterally in the horizontal segment of the intraparietal sulcus (or HIPS). Drawing on the results of a dozen independent EEGs, tomography, and fMRI studies, they argue that HIPS activation displays the features we would expect of a dedicated quantity module.

As one might predict, HIPS is more active when experimental subjects attend to quantities than when they attend to other stimulus characteristics. There is moreover evidence that thinking about the meaning of arithmetic concepts activates the intraparietal sulcus more than the mere sensory presentation of number stimuli. Thus, tasks involving number comparison, for example, result in significantly more HIPS activation than tasks involving the reading of number terms or numerals. This increase in activation is bilateral; both hemispheres of split-brain patients are able to compare number size. In normal, left-handed subjects however the right hemisphere HIPS shows a significantly larger increase in activation than the left. This is consistent with our earlier result that (at least some) semantic representations of quantity (or 'numerosity') are largely language-independent. In general, the level of HIPS activation is proportional to the amount of numerosity information a subject processes; HIPS is more active when experimental subjects compute two sums then when they compute just one. As well, interestingly, HIPS is more active when subjects subtract than when they multiply. (This last result is readily explained: most subjects store single-digit multiplication table results in rote memory; by contrast, they compute differences on the fly. If HIPS underwrites numerical manipulation, we would therefore

[121]Let me note that some experimental psychologists, notably Susan Carey, have gone well out of their way to maintain a dialogue with mathematicians and with philosophers of mathematics. My intent is to make a methodological point rather than a sweeping claim.

[122]There is electroencephalograph evidence Ravizzaa *et al.* [2008] for frontal lobe and cingulate gyrus involvement in some tasks. The cingulate in particular seems to be involved in error correction—an interesting fact given the previous chapter's emphasis on the importance of constraints.

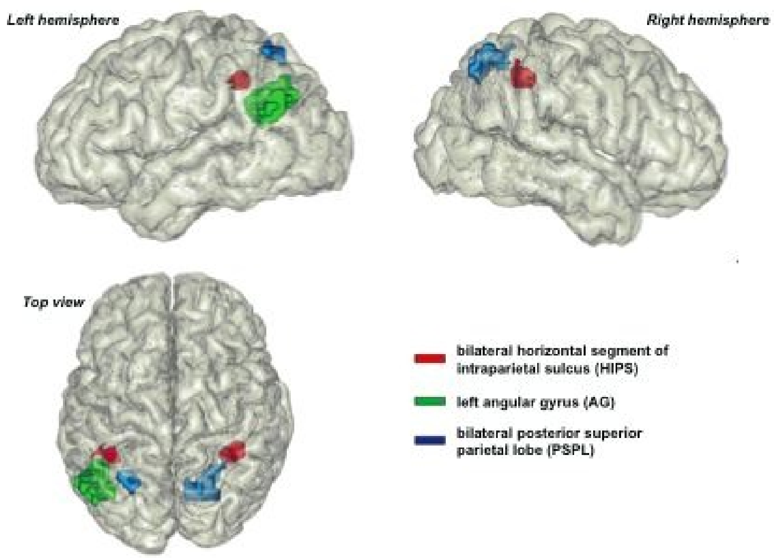


Figure 4: Cerebral localization of the mathematical cognition (from Dehaene et al. [2003]). Used with permission.

precisely expect it to be more active in subtraction than in simple multiplication.) Lastly, HIPS activation appears to be domain specific. Its activation does not appear to increase in experimental paradigms involving non-numerical comparison: for instance, the relative ferocity of animals, the orientation of objects, the position of body parts, and so on. That said, the absolute domain independence of the HIPS is still in question. Dehaene notes that further work is needed to establish whether HIPS is activated in comparison tasks which have a strong spatial or serial component (spatial prepositions, days of the week, months, etc.)

Not only is the horizontal segment of the intraparietal sulcus selectively activated by the processing of quantity, the magnitude and duration of its activation is affected by the magnitude of the quantities manipulated. This is noteworthy insofar is it may give us some clues concerning how numerosity is encoded. Neuronal activation is stronger and lasts longer when computing or estimating large quantities than small ones. In a comparison task, activation varies in direct proportion with the absolute distance between two numbers even when the numbers are presented as words. In the case of the integers, when no additional context is supplied—so when the implicit point of comparison is zero—activation is greater when the number is larger. The effect is robust, regardless of the medium in which the numerosity is presented: dot patterns, arabic numerals, spoken numbers and even spelled-out number names (T-H-R-E-E) induce the same effect. Nor does consciousness seem to play any discernible role. In one typical paradigm, a target number is presented. The target is preceded by a masked prime—flashed on a computer screen and subsequently overwritten too quickly to be registered consciously. HIPS activation occurs in all cases. The level of activation is higher when the difference between the prime and the target is greater (ONE vs. 4) as compared with cases when the target and prime are the same (1 and ONE). Again, the format of presentation does not appear to have any impact on these results; both numerals and words engender the effect.

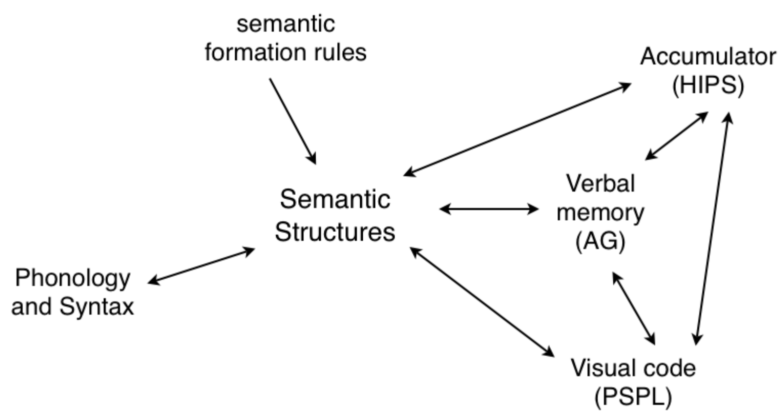


Figure 5: Cognitive processing at the language-math interface.

If HIPS were indeed the site of neuronal representations of numerosity, one would expect that the system could be selectively disabled, while leaving other mathematical abilities intact. This accords with what neurophysiological studies reveal. Calculation abilities are spared in (at least some) cases of severe semantic dementia resulting from a deterioration of temporal and frontal cortex. By contrast, almost complete loss of number-processing abilities following a small, localized lesion to the left parietal lobe has been reported [Cipolotti et al. 1991]. Interestingly, in that case, the capacity to understand numbers 1 through 4 was spared, suggesting a possibly distinct substrate for the processing of small numerosities (see below).

Accumulator. Several psychological models of the activity of a core quantity system have been proposed. Among them is the *accumulator*; a model advocated, among others, by

Gallistel, Gelman and Cordes [2005]. It's plausible that the quantity tracking system in the HIPS implements the accumulator or whatever psychological model of quantity estimation ultimately proves correct.[123]

Many animals possess a phylogenetically primitive system for representing continuous magnitudes. The existence of such a system has been demonstrated in rats, pigeons and a variety of other species [Dehaene 1997]. In a typical experiment a subject is presented with two levers, A and B. Depressing A at least some preset number of times has the effect of arming B so that subsequently pressing B releases a food reward. Pressing A too many times has no extra effect; however, pressing B early incurs a penalty. Many animals can be taught to perform the task even when A requires some two dozen activations before B is armed. The pattern of observed results is robust across species and quite striking. When averaged across trials, the number of arming activations gives a normal distribution with the mean slightly above the required number of presses. This indicates that animals are reasonably good as keeping track of numerosity information. The results should not be overstated however: response variability is a linear function of the number of arming presses required indicating that the subjects never learn to track precise number information. Rather, they seem to be responding to a rough estimate at each step.

Several aspects of the results are worth pointing out. It has now been shown experimentally that the pattern of responses is modality independent; the same results are obtained when sounds and flashing lights are used as stimuli or (remarkably) even when these are used in combination. This suggests that numerosity information is being represented by an amodal register of some sort (rather than, say, by the visual scratch pad or by auditory short term memory). Moreover, the number of presses an experimental subject makes is independent of its estimate of temporal duration; hungry animals press the levers faster but the same number of times [Dehaene 1997]. This reinforces the case for a dedicated system for representing numerosity.

Results from animals show that the capacity to estimate is widely distributed, hence phylogenetically quite primitive. Human beings, unsurprisingly perhaps, have been shown to display the same pattern of results. Once verbal code is overloaded, say by repeating a nonsense syllable rapidly, our capacity to estimate the number of lever presses we effect also (on average, over many trials) centers around the target number with a distribution proportional to the value of the number being estimated. (In other words, it displays scalar variability.) Two other features of these estimation results should be noted. It turns out

[123] Here, I set aside Dehaene's 2003 own preferred model, the logarithmic mental number line, in favour of Gallistel and Gelman's [2005] accumulator account. One reason is that the accumulator is more widely discussed in the literature; another is that the differences between the models are relatively minor—at least from a philosopher's perspective. My criticisms apply with equal justice to the accumulator and to its rivals. Note incidentally that the accumulator was used earlier to model temporal duration estimation [Meck and Church 1983].

that our ability perceptually to discriminate two cardinalities (as evidenced by reaction time data) decreases with the absolute value of the numbers; thus, 2 items are easier to tell from 5 than 72 are from 75. This is referred to as the *size effect.* Secondly, two cardinalities are easier to tell apart as the difference between them increases. This is the *distance effect.* (These effects are observed in adult humans, human children and nonhuman animals, reinforcing the case for a biologically primitive system.)

Gelman and Gallistel [2005] propose a cognitive model that purports to capture these phenomena. They point out that an object which stores a continuous quantity—say, a beaker holding liquid or an capacitor that stores electric potential—acts as a physical model of the observed results. They therefore propose that the brain employs a form of continuous quantity storage (or *accumulator*) to keep track of numerosity information. At each event being counted, the accumulator state is incremented by a fixed amount. Estimating the cardinality of a set of events or objects is a matter of 'reading off' the accumulator state. The reason, they suggest, that we find a normal distribution of results centred around the target value is the the process of reading information off of the accumulator is noisy. On average, we manage to get such readings right, but no individual trial is guaranteed to be correct. (Alternately, one could argue that the amount by which the accumulator is incremented at each step is itself noisy.) Finally, the size and distance effects are explained by on analogy with perceptual similarity. It's easier to tell apart two beakers containing 10 ml and 18 ml of water than to discriminate beakers containing 8 ml and 11 ml. The accumulator is, in sum, a model of *mental magnitudes* which in turn (it is claimed) are a representation system formally equivalent to the real numbers. Indeed, the suggestion is that the discovery of the real numbers in antiquity was (unbeknownst to classical Greek mathematicians) a sort of re-discovery:

> Our thesis is that this cultural creation of the real numbers was a Platonic rediscovery of the underlying non-verbal system of arithmetic reasoning. The cultural history of the number concept is the history of our learning to talk coherently about a system of reasoning with real numbers that predates our ability to talk, both phylogenetically and ontogenetically. [Gallistel et al. 2005]

I should mention as well that the accumulator, when yoked to the natural number system, purportedly affords us an understanding of the integers. That's because, the account goes, language consists of discrete symbols. These discrete symbols pick out the integers since integers are the numbers that represent discrete quantity [Gallistel et al. 2005]. (I will return to this topic momentarily.)

Left angular gyrus. The left angular gyrus houses a second cortical network implicated in arithmetic cognition. The AG is not a system dedicated exclusively to numerical cognition however. Rather, it plays a role in phoneme detection, reading, and in many tasks involving short term verbal memory. Performing certain calculations turns out to be among these. For instance, AG is more activated by precise calculation than by estimation tasks. It's also more active during multiplication than during subtraction tasks. And it's more active during tasks involving small products than large ones. Likely, this is due to a difference in how the information relevant to these tasks is stored. Relatively few subtraction results are stored in memory; by contrast, typically, many multiplication results—particularly single digit products—are memorized in advance. Results concerning addition further bear out this pattern; small sums (below 10) activate the AG, while larger ones, which must presumably typically be computed from scratch, do so less.

Bootstrapping. It's early days, so any proposed matching between AG and cognitive models must remain speculative. Still, it's hard not to view the AG as underlying one important component of a proposed representational synthesis which—according to an interesting hypothesis advanced by Susan Carey [2004]—helps give rise to children's understanding of the natural numbers. On Carey's account, we are 'bootstrapped' into an understanding of the natural numbers by learning to coordinate the activity of three cognitive modules: a 'subitizer', a sequencer, and the grammar of natural language itself (construed, here again, as an I-language).

Subitizing is the ability to register and attentionally track a small number of visually salient objects. It has been studied by Zenon Pylyshyn [2001, 1993] and his collaborators. Several factors distinguish subitizing from standard visual attention. First, the subitizing of up to three objects occurs very rapidly (perhaps as fast as 40-120 ms) while serial counting is much slower (250 ms or more per item). Second, objects registered via subitizing are picked out pre-conceptually. That is, they are not picked out by description or classified under a category. Instead, they are tracked purely on the basis of their spatial properties: their location and the pattern of their motion. Even relatively drastic changes in visual properties (colour, shape, and so on) do not disrupt subitizing. Moreover, it has been shown by Pylyshyn [2001] that subitized items can successfully be tracked while moving among visually identical non-target items and even while passing intermittently behind rigid occluders. The interpretation of subitizing data is still contentious. However, one way to account for it is to think of relatively early visual processes as registering each of the three objects in a separate 'object file' and subsequently keeping track of it by location and trajectory.[124]

Clearly, subitizing alone (or even in conjunction with accumulator data) is incapable of giving us knowledge of the natural numbers. Any object files created by the process are unconnected to one another and stand in no formal relations. They therefore fail to represent

[124]Subitizing may not be an early visual process at all. It may be related to the 'where' path in Goodale's 1995 two-path account. The issue is not crucial for us here.

such basic facts as that each number has a unique successor and that every number (but zero) is a successor of some number. This is consistent with the experimental evidence: infants already possess a capacity to discriminate gross differences in numerosity, such as that between one and two stimuli [Wynn 1992, 1998]. However, an understanding of the meaning of number terms, such as "one", "two", or "three", is a hard-won ability that develops gradually between the ages of two and four. How then is such knowledge acquired?

Carey [2004] argues that basic numerical understanding depends in part on an understanding of language.[125] Around the age of two, children appear to treat small number words on a par with quantifiers, such as 'many' and 'some.' They understand 'one' to refer to collections of a precise cardinality. By contrast they initially take 'two' and 'three' to be interchangeable with 'many.' Thus, asked for three toys, young children will hand over a handful. The result is robust across language communities and cultures. Over the space of several months, children come to grasp that 'two' designates only sets of a determinate cardinality. An understanding of 'three' as referring to yet another cardinality follows some months after that. The mean age at which children become two-knowers and three-knowers varies between language communities. In particular, languages that have explicit morphological markings for duals in addition to plurals (Russian) facilitate acquisition; languages that do not mark plurals (Japanese) slow the process slightly. In any case however, all neurotypical children arrive at the same end-state.

Three-knowers do not yet have an understanding of natural numbers or integers *per se*. Notice though how far ahead they are of even the most capable chimpanzees in terms of their arithmetic comprehension. Chimps can be taught to associate numerosities with printed, Arabic characters. It has been shown that numerals between one and nine as well as zero can be learned. As well, chimps can associate such symbols with the cardinalities of collections of token items [Kawai and Matsuzawa 2000]. And they can perform arithmetic operations, including addition and subtraction using the symbols alone or using a mix of symbols and token items (oranges, say). There are however important disanalogies between chimpanzee knowledge of arithmetic and ours. Their performance remains error-prone even at small cardinal values; they never cease to occasionally mistake 5 for 6, or 7 for 8. This suggests that they may continue to rely on approximation (perhaps mediated by the accumulator) rather than grasping a truly discrete counting system. Also, it has been noted by Matsuzawa and his collaborators that chimp performance with the symbols degrades slightly after each new symbol is introduced. Again, this result suggests that a somewhat different underlying cognitive process than ours underlies chimp arithmetic. Human children who learn what 'three' means continue to understand 'one' and 'two' perfectly well. Most importantly

[125] I will make no attempt to reconcile this hypothesis with the evidence presented in Chapter 2. These are open empirical issues. Please note that the final argument against Platonism presented in that chapter does not depend on how this controversy turns out.

though, chimps never grasp the successor function and so do not extrapolate from small number symbols to the existence of an infinite set of naturals.

> [C]himpanzes understanding of the integers is based on a mechanism that is wildly different from that of human children. Chimpanzees learn the integer list by brute association, mapping each symbol to a discrete quantity. Human children, in contrast, learn by making an induction from a limited body of evidence. Children induce that the integer list is created by a successor function, and that this function generates an infinite list of numbers. [Hauser and Spelke 2004]

In fact, it's unclear how human children take that important next step. What seems to happen is that three-knowers next induce not just the number four but also the fact that each natural number has a unique successor. Natural number quantifiers and the visual subitizer are not sufficient to accomplish this. Carey [2004] suggests that the final necessary component of the human counting system is the capacity for representing serial order. Children learn to recite various ordered lists relatively early on and typically with no understanding of the meaning of the terms. Think here of the alphabet song, counting rhymes, or skipping-rope songs. Carey's suggestion is that kids learn the counting sequence by rote, storing it in verbal memory (perhaps mediated by the AG). Subsequently, as children come to realize that the first three items in that sequence not only designate an item on a list but also stand for sets of objects that differ by a cardinality of one, they make the crucial step. They come to realize that *any* word in the number sequence designates a cardinality. "For any word on the list whose quantificational meaning is known, the next word on the list refers to a set with another individual added" [Carey 2004]. The claim is that here we find the cognitive basis of our understanding of the natural numbers and (later) of the integers. (Note, incidentally, that Carey's theory makes no essential use of analogue magnitudes or the accumulator. It is therefore substantially different from the model advanced by Gallistel et al. [2005].)

Posterior superior parietal. Let's return to the neuroscientific evidence once more. The third, and perhaps most enigmatic parietal system involved in mathematical processing is the bilateral posterior superior parietal lobule, or PSPL. Like the AG (and unlike the HIPS) this system is not domain specific. It plays a role in a number of visual and spatial tasks, including hand reaching, grasping, and eye-orienting. It is activated in mental rotation tasks. And it is involved in spatial working memory. It can be thought of as implementing some of the functions of the psychologists' visual-spatial scratch-pad. PSPL is activated in number comparison, approximation and subtraction tasks. The hypothesis advanced by Dehaene [2003] is that the PSPL is an attentional system involved in sequentially focusing on particular segments of the internalized representation of the real number line. The hypothesis is, clearly, somewhat speculative. In any case though, the mandatory involvement of a spatial attention region in calculation (and perhaps other mathematical tasks as well) is suggestive and we'd do well to bear it in mind.

Limitations

I don't think there can be much doubt that ongoing research in neuroscience and in cognitive psychology is making important strides toward an understanding of the systems that underlie human mathematical abilities. In all likelihood, a good deal of the work being done today will eventually constitute a part of the finished theory of our mathematics faculty. Nonetheless, as I have already intimated, the models still have significant limitations. Let me explain what I take those limitations to consist in before offering some suggestions for how we can move forward.[126]

Theory. Let me begin with some general methodological considerations. A fully elaborated theory in cognitive science is pitched at three distinct yet complementary levels, the *Marr hierarchy* (after Marr [1982]). The *highest* level is that of *task description.* Here, we offer an characterization of the cognitive operation being performed that specifies what the agent is trying to accomplish. The characterization is intentional, in Dennett's [1987] sense; it avers freely to what the subject believes, hopes, fears or intends. Some examples of task descriptions include: recalling a numeral from memory, attending to a blue dot, distinguishing two pictures, listening to an auditory stream presented to the left ear, and so on. By contrast, the *lowest* level of description offers an *implementational* account of the task—that is, a description in physiological terms. Theories pitched at this level are the domain of the cognitive neuroscientist. A great deal of the action in cognitive science takes place at the crucial middle layer where the (often implicit) aim is to bridge the gap between task description and implementation. This is the domain of the cognitive psychologist, linguist, and computer modeler. Research at this level is predicated on two (by now familiar) assumptions: that the mind operates by manipulating information; and that it does so by computing, in the sense of Turing [1950]. It may well be that the information is distributed over vast, structured neural networks and that the computation itself is massively parallel.

[126]Before offering some criticisms, let me reiterate my motivations. Regrettably, philosophical objections to psychological research concerning mathematical reasoning have in the past sometimes been dismissive and unhelpful. Some argued, for instance, that since the referents of mathematical terms—mathematical objects—necessarily transcend any attempt to make sense of them in terms of pedestrian empirical posits, psychology and physiology have *nothing* of value to contribute to an understanding of mathematics *per se* (see especially Frege [1953]). It should now be apparent that I want no truck with that view. In any case, it seems slowly to be in decline. Martin Kusch [1995] has recently written a fascinating history of early 20^{th} century anti-psychologism. A number of respected philosophers have, moreover, started to back away from more radical variants of the doctrine. See for instance Thomason [1995], Haack [1978], Giaquinto [2001], and Laurence and Margolis [2005].

Even so, the tri-level description holds.[127]

Two sorts of theories at the all-important middle-layer are possible. The one sort is an account of the interfaces and the flow of control between coarse-grained mental functions. This is sometimes called a theory of processing (or a boxology). The other sort of theory is an account of the manipulated representational states themselves: a theory of structure.[128] A cognitive model is said to be *weakly equivalent* to human mental performance if it mimics it sufficiently well so as to be input-output equivalent to us. Computer programs that play chess (and include time-delay and error-production subroutines) are a good example of weakly equivalent models. A piece of C^{++} code based on standard data-structures which proved capable of passing the Turing test would be weakly equivalent to human beings as well; the mind/brain doesn't rely on linked lists on which standard data structures ultimately rely. In sum, weak equivalence can theoretically be achieved by sufficiently sophisticated models of processing. By contrast, a cognitive model is *strongly equivalent* to a human in some cognitive domain if it performs a mental task by means of an algorithm that manipulates structured representations identical to the ones in fact used by the mind/brain. The ultimate aim of cognitive science is to offer strongly equivalent models of human cognition. (For a discussion of the tri-level approach to cognitive science, see Dawson [1999], Dennett [1998] and Pylyshyn [1984].) Let me now suggest that this characterization of the cognitive enterprise helps shed light on what we ought to be looking for in an explanation of our mathematical abilities. What we want are cognitive models that perform the same tasks that human beings perform and do so by the same means. Unfortunately, it seems on reflection that current models fall short of this goal.

Task description. The lack of strong equivalence between cognitive models and human minds manifests itself in a variety of ways. The most basic is that some of our models appear to be executing functions that, contrary to the advertising, are not formally equivalent to those being performed by the lay mathematician.

Reals. Gallistel and Gelman are careful not to *identify* mentally represented magnitudes with real numbers, preferring to leave such ontological questions alone. They do however claim that mental magnitudes are "formally equivalent to the real numbers" [Gallistel et al. 2005]. Given the semantic theory we adopted in the previous chapter, this is nearly as strong a claim. If we had at our disposal a thorough account of a mental module whose representations were truly formally equivalent to some mathematical structure, we would be a short step away from a theory of how our semantic representations of that mathematical

[127]See Dawson [1999] for an extended discussion of the tri-level hypothesis in a connectionist setting.

[128]A clear distinction between theories of structure and theories of processing is made, for instance, by Jackendoff [1987]. Generative grammar is one example of a theory of structure. Biederman's [1995] theory of visual object recognition is another.

structure were generated and constrained. So if Gallistel and Gelman are right about the formal equivalence, we are closing in on a cognitive model of human ability.

On closer examination, it looks however that states of the accumulator are not formally equivalent to the real numbers after all. The first problem is that the reals are dense; between any two real numbers there exist an infinite number of other reals. It's implausible that accumulator states are densely packed since this would require the accumulator to encode information with infinite precision. It's hard to see how the human brain could be capable of doing that. And there is no evidence that rat, pigeon, or chimp brains do so either. (In fact, things are worse still. The reals form a continuum; the cardinality of the set of real numbers is $2^{\aleph_0}$.) Secondly, the reals are unbounded upward; there is no greatest real number. Again, it's hard to see how accumulator states could be similarly unbounded since they are implemented in a physical system.[129]

A third difference runs in the other direction. Accumulator states have successors since—according to the model being proposed—we increment them while estimating quantity. The accumulator is hypothesized to include a computational (effective) procedure for generating the successor of each state.

> Importantly, the verbal counting process is homomorphic to the nonverbal counting process. In particular, both processes have effective procedures for defining successor symbols. Each step in the verbal process summons the next word from the list of count words. Each count in the nonverbal process defines the next magnitude [Gallistel et al. 2005].

As we know, there is no effective procedure for determining the successor of a real number since no real number (*qua* real number) has a successor.[130] Formal equivalence is a transitive relation. If A is formally equivalent to B and B to C then it follows that A and C are formally equivalent also. If accumulator states are homomorphic with the natural numbers (as, it seems they are) and the natural numbers are not homomorphic with the reals then, it follows, accumulator states are not homomorphic with the reals. But if so then we still lack a model strongly equivalent to human arithmetic ability.

Admittedly, one might challenge this line of reasoning by saying that, while counting, the accumulator increments its state by *roughly* one (construed as a real). Thus, it increments by

[129] Dehaene's algorithmic number line model does better in this regard. It however requires that the amount by which the counter is incremented decrease exponentially and with infinite precision.

[130] There's a wrinkle here. Recall that on the structuralist account (which we did not reject) mathematical entities are characterized entirely in terms of their relations to other members of the structures which they inhabit. Structuralism permits us side-step potentially embarrassing questions about just *which* sets are the *true* surrogates of the natural numbers. Questions about relations within structures can be posed and answered meaningfully; matching elements across structures however is ultimately a matter of convenience. It is ultimately done by fiat. The number 2 construed as an integer, of course, has a successor. The real number that is customarily matched with it does not.

1.0001 on one occasion, by 1.03 on another, and so on. This however does not help. Suppose we admit this and suppose we also try to follow Gallistel and Gelman in trying to explicate the arithmetic relations in terms of accumulator states. Doing so, we are led to define *equality* of accumulator states as something like *rough similarity.* That's because we have abandoned the notion that some accumulator state n, incremented m times on two different occasions will yield precisely the same result. The lack of a precise equality relation between accumulator states seems to differentiate them importantly from real numbers. After all, $\pi + \sqrt{2} + 1$ computed on two different occasions yields exactly the same answer. It's hard to see therefore how accumulator states and real numbers can truly be formally equivalent. And so we still seem to lack an appropriate model of human arithmetic ability.[131] But if there is no formal equivalence then Gallistel and Gelman's claim that in discovering the reals, the ancients made a platonic re-discovery of the accumulator and its states looks implausible.

Integers. There is another problem with Gelman and Gallistel's [2005] theory. Even if the accumulator account could be shown to model a complete ordered field and to underwrite human ratiocination concerning the reals, we would still lack a convincing account of the integers. Recall that Gelman and Gallistel attempt to derive the integers from the coordination of activity between accumulator and natural language.

> When a discrete system like language attempts to represent quantity, it will find it much easier to represent countable (discrete) quantity than to represent uncountable (continuous) quantity... [T]he integers are picked out by language because they are the magnitudes that represent quantity. Countable quantity is the only kind of quantity that can readily be represented by a system founded on discrete symbols, as language is. [Gallistel et al. 2005]

The claim made in this passage is puzzling. It seems clear that there is no direct dependence between representational medium and the domain represented. An analogue model can represent discrete entities, just as a digital model can (perhaps imperfectly) represent a continuous magnitude. Natural languages have the resources to represent both. We have words like 'dozen' as well as 'twelve'. We can keep track of how many chocolates are in the box just as easily as we can keep track of how much chocolate is there. But let's be charitable. Fundamentally, the problem with the model (even setting aside the various claims about natural language) is that it does not offer the cognitive modeller much to go on. It's hard to understand precisely *how* to yoke a capacitor circuit to a computational model of the grammar for a natural language so as to obtain a representational system capable of manipulating the natural numbers. And without those sorts of details, it's hard to assess whether the model is plausible.

[131] For related though slightly different criticisms of Gallistel and Gelman's model, see Laurence and Margolis [2005].

By contrast, Susan Carey's two-interface model of the acquisition of natural number concepts is both elegant and convincing. It's possible that Carey's work offers us an insight into one aspect of the functioning of the mathematics faculty. But even if that's right, ongoing work in cognitive science still faces two other sorts of philosophical objections.

Scope. An important symptom of the inadequacy of our models *qua* theories of structure concerns their scope. Nearly all current work—including book-length treatments by Butterworth [1999], Campbell [2005], Dehaene [1997]—discuss grade-school arithmetic to the exclusion of all other branches of mathematics. We have no models of human set-theoretic, group-theoretic, or topological reasoning, for instance.[132]

Two replies can be made. One might argue that the cognitive operations involved in higher mathematics involve the same underlying operations that are used in basic arithmetic. Unfortunately, there is little psychological evidence for this supposition. We are being asked to believe, in effect, that cognitive processes that range over basic arithmetic information are sufficient to allow one to represent and track the full range of mathematical facts. However, the fact that set theory—rather than, say, the Peano axioms—is typically taken to constitute a suitable foundation for mathematics makes this reply highly implausible. The second way to respond is to dismiss higher mathematics as essentially a cultural invention.[133] I don't think this sort of deflationism can be made to work. In the opening chapter, I offered reasons for supposing that a good deal of higher mathematics is contentful, even when we are initially disposed to suppose that it is not. The arguments advanced there hinged on higher maths' uncanny utility in providing a conceptual scaffolding for original scientific research in the natural sciences. A psychologist tempted to dismiss group theory or non-Euclidean geometry as empty sophistication owes an account of maths' indispensability. In the absence of an account, we have little reason to abandon the minimal realism defended earlier. It seems therefore fair to conclude that ongoing cognitive investigations leave out vast swaths of mathematics.

Even if the point is correct, it may seem insignificant. All scientific investigations are forced by the nature of the enterprise to limit their purview. In this case however, I think that the limited scope of investigation not only artificially constrains investigation; I think it distorts results and leads to theories that will not scale up. To see why, consider an analogy with linguistics. Imagine for a moment that generative grammar had not been developed. Imagine that psycho- and neurolinguists studying language had chosen to focus (perhaps understandably) on four 'basic' linguistic operations: asserting, asking questions, making requests, and issuing warnings. There is little doubt that stable correlations could

[132] Recently, a few researchers, notably Lee and Spelke [2008], have begun to investigate Euclidean geometry in addition to basic arithmetic. Even so, cognitive science is still far from embracing a broad realism about the mathematical universe.

[133] At least one well-known psychologist has advocated this view.

be found between cortical activation and the emitting of assertive, interrogatory, soliciting and cautionary linguistic behaviours. Interesting comparisons can undoubtedly be found between those cortical activation patterns in humans and in nonhuman animals. Moreover, I have little doubt that matching boxologies—i.e., theories of processing—could be drawn up to 'account for' the results. Unfortunately, all that work would get us no closer to proper neurolingusitics in the modern sense [Pulvermüller 2002]. The problem is that a theory of speech acts, no matter how nuanced, is no substitute for generative grammar. And it's largely the latter and not the former that helps inform the neuroscientist what she ought to be looking for in patterns of cortical activity. It's for this reason, I believe, that psychologists who are content to operationally define 'mathematics' as little more than the four arithmetic operations are committing a significant theoretical error.

Philosophical desiderata. It's perhaps slightly unfair to expect psychological models to meet metaphysical desiderata. Fair or not though, the fact remains that available models of arithmetic reasoning fail to account for several aspects of mathematical knowledge that philosophers and philosophically-minded mathematicians care about.[134]

The first such philosophical problem concerns maths' much touted *necessity*. To put it picturesquely, the very same mathematical objects with the very same properties exist in any possible world. There is no way that the world could have been such that the mathematical entities, properties and relations had been different. The trouble with cashing out such entities, properties and relations in terms of the states of psychological models (including the accumulator or other sub-faculties) is that psychological and physiological facts don't go that deep [Frege 1953]. Our psychology could have been different. In fact, it can be altered: there is evidence that accumulator performance can be affected by the administration of amphetamines; under such conditions the subjects overestimate cardinalities [Dehaene 1997]. Surely though the real numbers themselves don't change under those circumstances. Evidently, we need an independent standard of correctness, one not defined in relation to contingent facts. Here's another way of making the same point: it's possible that somewhere in the cosmos there exist intelligent beings with an altogether different psychology and physiology. It's possible, moreover, that some of those beings have developed natural sciences that permit them to study their surroundings, overcome their folk prejudices, and engineer useful tools. Surely we'd want to concede that they have an algebra, geometry, and calculus. And so we need some way of characterizing the objects and relations that constitute those which itself does not spell out the relevant posits in terms inextricably linked to avian and mammalian physiology.[135]

[134] See Jerzykiewicz and Scott [2003] for a related discussion.

[135] I'm not sure a move to functionalism helps us in this case unless we can also find some way of picking out the relevant functions by means other than mathematical notation itself. I leave this as an open question.

The second property of the mathematical universe that needs explaining is the *applicability* of mathematics in natural science itself. Mathematical (or purely formal) descriptions of natural phenomena apparently have more power than ones offered in natural language. I addressed the issue at length in Chapter 1 and I have nothing new to add here. My point simply is that a philosophically satisfying account of mathematics should explain this fact. Proposing that an accumulator and a subitizer (and so forth) interface with semantic representations offered up by the language faculty does nothing to dispel the air of mystery. It therefore falls short of a philosophically satisfying account.

Mathematical 'perception'

Eventually, we want a fully elaborated tri-level account of human mathematical abilities (in the style of David Marr [1982]). Evidently, that will require a theory of the nature of the 'representations' employed in mathematical reasoning. It's perhaps not surprising that I have no concrete theory to offer here. Instead, I'd like to suggest a way of understanding the nature of mathematical cognition that may help us arrive at a theory. The route will be slightly circuitous and will necessitate a certain amount of historical exegesis. In this case, that can't be helped (and in any case, I think the history is intrinsically interesting).

The first step will be to abandon the first-pass, commonsensical understanding of 'mathematics' as consisting primarily of reckoning; we'll need to dig a bit deeper. In Chapter 1 we encountered Alain Connes' bold assertion that in the course of his research, he senses himself to be charting an abstract landscape every bit as real as our physical surroundings. This metaphor of mathematics as a kind of perceptual experience can, of course, be traced back to Plato. I'd like to suggest that it is precisely to *this* conception of the nature of mathematics that cognitive science ought to pay heed if we hope ultimately to make sense of mathematics as a natural phenomenon. I'm not deaf to how odd that sounds, nor am I about to betray my naturalist scruples. Let me offer the hard-nosed empirical psychologist a reason not to abandon the discussion at this stage. Consider: as it happens, few adults are synesthetes. To neurotypical subjects descriptions of the experiences of synesthetes sound utterly fantastic and wholly implausible. On its face, little if any sense can be made of such assertions as that peanut butter is full of circles and spheres or that high-pitched sounds have an unpleasant briny taste.[136] In spite of this, in order to arrive at a scientific account of synesthesia, a good first step was to treat the subjects' reports concerning their phenomenal experiences perfectly seriously. Indeed, such reports have become a respectable point of departure for experimental investigation. We are now able to match the phenomenological descriptions supplied by synesthetes with correlated psychological and neurophysiological

[136] The examples are from Cytowic [2003].

events.[137] Think now of the mathematicians' inner experience. Given the profound importance of their field to our cultural and scientific endeavors, I see no obvious reason to deny them the same courtesy. As in the case of the synesthete, it's not a matter of taking their self-reports as authoritative. Rather, it's a matter of making these the starting point for investigation. Anyway, we have little to lose by provisionally taking seriously the mathematician's descriptions of her phenomenology as well as her attempts to offer a philosophical interpretation of them.

The modern *locus classicus* of the mathematicians' sophisticated self-analysis is a short [1947] essay by Kurt Gödel entitled *'What is Cantor's continuum problem?'* There, Gödel famously asserts that

> Despite their remoteness from sense experience, we do have something like a perception of the objects of [fundamental mathematics, and in particular] set theory.

Gödel is an appropriate target for detailed study. His voluminous correspondence, notes, and essays afford us a wealth of material. Moreover, Gödel was a careful and penetrating thinker. We know a good deal about his intellectual biography as well as his metaphysical and mathematical commitments and so it's possible to reconstruct what he intended by his claims in considerable detail. Once the interpretation is complete, we can begin to translate the results into a more familiar, modern, cognitive-scientific idiom. I will argue toward the end of the chapter that some progress toward an epistemology of mathematical judgements can be made in this way. Gödel's work also suggests how the structures that mathematical reasoning manipulates should be understood.

Background

The key to understanding Gödel is to understand Kant. This, I realize, is not the received wisdom. Indeed, Gödel has been called an arch-Platonist by more than a few commentators.[138] Before pressing ahead therefore let me briefly justify my departure from the standard view.

We know that Gödel become a 'conceptual and mathematical realist' in 1925: early on in his undergraduate career. The conversion coincided with his second, sustained attempt to study Kant's works. The first foray had already been made in 1922 while Gödel was still a teenager attending *Gymnasium* in Brno. At that point, the young mathematician had

[137] See also Dennett's [1991] discussion of *hererophenomenology.*

[138] This claim is made *inter alia* by Brown [1999], Shapiro [1997], Steiner [1998], as well as his biographer Wang [1987]. In fact, even some recent commentators who recognize the Kantian roots of Gödel's ideas try to reconcile those with a Platonist ontology (Martin Solomon [2009]). Gödel is undoubtedly a realist, in the sense discussed in Chapter 1. That he is a Platonist, in the sense addressed in Chapter 2 would require considerable further evidence.

mastered undergraduate mathematics and was beginning to turn his attention to related philosophical matters.[139] By 1925 Gödel was sufficiently taken with philosophy to enroll in Gomperz full-year class on the history of the field. During that year, he read *The Metaphysical Foundations of Natural Science*, evidently as background for a reading of the first *Critique*. The impact of Kant's work on Gödel cannot be overstated. Indeed, it would not be an exaggeration to say that Kant constituted *the* major intellectual point of reference for Gödel throughout his life.[140] Of course there were others. We know that several other philosophers, notably Husserl and Leibniz, also exercised an influence on Gödel later on in life. Yet in 1975, when asked to reflect on his intellectual trajectory and to name the philosophers who had contributed to the *development* of his ideas, Gödel responded that "only Kant was important."[141]

Let me temper what I have just said with a word of caution. The question of intellectual influence in this case needs to be treated delicately. Gödel was an original thinker in his own right, so while Kant exercised a powerful pull on him, it would be a serious mistake to think of Gödel as a 'Kantian.' In a 1961 essay, he writes: "a general feature of Kant's assertions [is] that literally understood, they are false, but in a broader sense contain deeper truths." The first challenge for any interpreter of Gödel's philosophy must therefore be to disentangle the views he actually held from the rich Kantian backdrop against which they are defined. In order to do this, it's crucial to understand where Gödel departs from transcendental idealism as well as his reasons for doing so.

In what follows, I propose to pursue these issues. Since the aim, ultimately, is to be in a position to interpret Gödel's [1947] discussion of the nature of our access to mathematical reality, I will be particularly interested is relatively early critical philosophical reflections on mathematics and metaphysics—those written between 1946 and 1949. As it happens, we are in luck—there is concrete evidence to be had. The [1947] paper, entitled *What is Cantor's Continuum Problem?* was commissioned in February 1946 by *American Mathematical Monthly*. Gödel took the writing of this article—which had apparently been intended by

[139]See [Dawson 1997].

[140]For an interesting historical discussion see Yourgrau [2005].

[141]The questionnaire can be found in Gödel's *Collected Works*, volume IV as well as in [Wang 1987]. Oddly enough, Wang [1987] recognizes that in 1925 Gödel read a good deal of Kant and that he became a 'conceptual realist' in that year. Yet he inexplicably suggests (p.22) that the conversion to realism occurred "presumably under the influence of Plato." He offers no evidence for this presumption.

I should mention that some interpreters—notably Charles Parsons [1983]—suggest that it may have been Husserl and not Kant that Gödel has in mind in his 1947 essay. This is certainly an interesting idea and the connection between Husserl's work and Gödel's later philosophy is worth pursuing. Still, I find this reading unconvincing. Gödel did not devote himself to a serious study of Husserl's philosophy until 1959. He had became slightly familiar with Husserl's work, especially work on consciousness, when the latter gave two lectures in Vienna in 1935. Still, even bearing this in mind, I think connecting the 1947 paper to Husserl would require additional evidence. In fact, Leibniz, whom we know Gödel studied between 1943 and 1946, and maintained obsessed with throughout his life, might be a more plausible alternative candidate.

the journal's editor to be a relatively straightforward expository piece—very seriously. In fact, he became quite absorbed in the project. Ultimately, he required an extension. And the paper was not sent to the journal in its final state until May 29, 1947. While in the midst of writing the piece on Cantor—in May 1946, to be precise—Gödel was asked by Paul Schilpp to contribute to a volume in honour of Einstein, to be given to the latter on the occasion of his seventieth birthday (March 14, 1949).[142] Of course, Gödel accepted and so took on a second project, which added significantly to his workload. Shortly thereafter, in September 1946 he wrote to his mother that he had become "so deeply involved in his work that he found it difficult to write letters" [Dawson 1997]. This second project gave rise to several unpublished manuscripts entitled *Some observations about the relationship between theory of relativity and Kantian philosophy*. It also resulted in an abridged piece that was ultimately given to Einstein. Now, what makes these papers particularly valuable from our perspective is that they contain an extended discussion and critical assessment of Kantian metaphysics from a modern standpoint *written at precisely the same time* as the [1947] remarks on mathematical perception. Indeed, since Gödel sent off the paper in which the above cited passage is contained in May 1947, drafts of the Kant essays that were written just before that point are of particular interest. Here again, we are in luck. We know that Gödel made an important technical discovery on September 23, 1947.[143] I propose here therefore to work mainly with a draft of the Kant essay written immediately before that.[144] My contention is that the metaphysical picture Gödel lays out in that paper is essentially the same as the one that serves as the background for the [1947] essay.[145]

Kant's Transcendental Psychology

Although Gödel is not an orthodox transcendental idealist, he does argue that much of the Kantian world-view is coherent and broadly correct in its essentials. He does this by showing that modern physics—and the general theory of relativity in particular—offers a striking confirmation of some of the core metaphysical tenants advanced in Kant's [1781] first *Critique*.[146] Simultaneously, Gödel [1949] takes the opportunity to clarify precisely

[142]The paper was, in fact, ultimately not included in the volume but rather given to Einstein directly at his birthday celebration, on March 19, 1949.

[143]He discovered that relativity permitted worlds in which a simultaneity relation cannot be defined. See the discussion by Malament elaborated by Stein in Gödel's *Collected Works*, volume III, pp.203-4.

[144]This is the B2 draft written between the fall of 1946 and September 1947.

[145]The alternative is that while working intensely on three conceptually difficult topics—the continuum, general relativity and transcendental idealism—Gödel nonetheless entertained some *other* metaphysics and left no traces of this in the written record. I find such an alternative hard to accept. The more so, in fact, since the period I am focusing on antedates Gödels interest in Husserlian phenomenology.

[146]Note that I read the *Critique* as mainly an ontological and not as an epistemological treatise. In fact, I understand it as a contribution to transcendental psychology in the way suggested by Kitcher [1990]. My reading is also heavily indebted to the interpretive work of Brook [1994], Carson [2004], Falkenstein [1995].

which aspects of transcendental idealism have been superseded by subsequent developments in mathematics and natural science.

Faculties. Central to the *Critique* is a distinction between two cognitive faculties—sensibility and intellect—as well as their complex interplay. With the benefit of hindsight, it's easy to see Kant as a revolutionary. So it's helpful to bear in mind that at first glance the bicameral architecture would have struck Kant's contemporaries as a rather reactionary commitment.[147] Both rationalists and empiricists had favoured single-facultly accounts of the mind. These however, in Kant's estimation, had run their course and failed: "Leibniz intellectualized appearances, just as Locke had sensualized all of the concepts of the intellect, i.e., had passed them off as nothing but empirical" [A271]. By returning to an earlier theory—one introduced by Aristotle and employed by the scholastics—Kant effectively turns back the clock. In his hands however, the old *aesthesis/noûs* distinction takes on new life. For reasons that will emerge shortly, the faculties are characterized by Kant exclusively in terms of the *functions* they perform rather than (say) in terms of the physiological systems that implement them. This articulation of early functionalism is already a major original contribution to the philosophy of mind [Brook 2004]. Just as important is the tightly constrained range of functions that Kant is prepared to ascribe to the two faculties.[148]

Kant denies that we can have direct insight into the workings of the mind via clear and distinct ideas or by any similar, privileged means. Nevertheless, we can infer a good deal about our constitution by attending to the mind's products—our experiences—and inferring the structure of the mental faculties that contribute to producing them.[149] One striking aspect of our conscious experience is the lack of control we have over its basic content. We cannot alter the state of our surroundings by the application of willpower alone. Facts about our surroundings 'force themselves upon us as being true.' By contrast, our dreams and our imaginary phantasms, in this respect at least, are largely under our control. To account for this, Kant suggests we must posit the existence of a receptive mental faculty, the *sensibility*, and allow that it receives input from without. Since he claims no privileged insight into its nature, his characterization of the sensibility is appropriately minimalist: it is construed as engaging in no inference-making, no reckoning, and as drawing no conclusions. Instead, the

I think Gödel's claims concerning the affinity between Kant's and Einstein's work bears out this reading. Einstein, after all, was not offering an epistemology; he was discussing what there is.

[147]For a discussion, see Falkenstein [1995].

[148]Let me note in passing that the bicameral architecture permits Kant to resolve some outstanding problems, not the least of which is the dispute between Leibniz and Newton concerning the nature of space. Discussing this and the distinction of the faculties in Kant's earlier work, such as the *Inaugural Dissertation*, would take us too far off track. It is however independently interesting to anyone curious about the ur-history of cognitive science. I refer the interested reader to Carson [2004] and Falkenstein [1995].

[149]This sort of *transcendental argument* is now commonplace in linguistics. We look to speaker performance and native speaker judgements so as to hypothesize a theory of the grammar that produced them. See Chomsky [2000].

sensibility merely places us in an *immediate* and synchronic contact with that which forces itself upon us: viz., concrete particulars.[150] These imprint themselves upon us, thereby becoming sensory impressions. True to form, Kant makes no special claims concerning the distinction or the relationship between auditory, tactile, olfactory, and visual impressions. Questions about the individual senses, their interrelations and the sensory organs themselves are left to empirical psychology and physiology. Given how little we can safely infer about the functional characterization of the sensibility from transcendental arguments based on the nature of phenomenal experience, this is surely the prudent approach. (It prevented Kant from making rash commitments that would have been disconfirmed by Galton's work on synesthesia, for example.) Instead, Kant uses the neutral term 'intuition' (*intuitus*) to refer to the generic capacity to receive imprints or impressions (*Eindrucke*). It's worth underscoring that this term picks out a wholly pedestrian, non-mystical process: "Our *intuition*, by our very nature, can never be other than *sensible* intuition; i.e., it contains only the way in which we are affected by objects."[A51] Kant therefore rejects all manner of extra-sensory or supernatural input to the mind. (Gödel does exactly the same in the [1947, 1949] papers we will discuss).[151]

Human beings may have a tendency to overestimate our species' cognitive prowess and devalue that of others. Still, we surely do enjoy a modest measure of mental flexibility and a capacity for spontaneous thought. This too requires an account. Kant attributes this aspect of our experience to the second part of the bicameral architecture, the *intellect*. The powers of this active faculty are essentially discursive. It proceeds by drawing (largely unconscious) inferences and arriving at judgements. To do so, the intellect subsumes the sensory impressions delivered over by the sensibility under concepts. Unlike the synchronic and unmediated action of the sensibility, the activities of the intellect unfold diachronically; they take time to accomplish. And, of course, they can be executed incorrectly; we sometimes subsume a sensation under a concept only to discover on further reflection or in light of subsequent experience that the predicate we had employed does not truly apply. Think here of erroneous face recognition. Or again, think of visual illusions.

According to the mature Kant of the first *Critique*, all representations manipulated by the intellect, without exception, derive their content mediately or immediately from the sensibility [A51]. The intellect contains no innate ideas, exactly as the empiricists had maintained.[152] Nor does Kant ascribe any special receptive powers to the discursive faculties

[150]Zenon Pylyshyn's [2001] *visual indexes* are a direct modern extension of Kant's doctrine of sensibility.

[151]Interpreters of Gödel, including Shapiro [2000], often take 'intuition' to mean a sort of hunch or gut feeling that some proposition must be true. This is a misinterpretation.

[152]Kant does not however deny that the mind contains innate 'grounds' of representation. There is some controversy over whether these should be seen as innate mechanisms or something more basic, more passive than that (see Kitcher [1990], pp.16-17). On the innate mechanisms reading, Kant can be read as taking up Hume's [1748] challenge to 'characterize and class' the principles by means of which the association of ideas

of the mind. This is a significant limitation. It directly contradicts the scholastic doctrine that *noûs* permits us to directly grasp certain 'higher' truths (a thesis Kant had tolerated in earlier work). Kant's mature transcendental psychology offers us an austere vision of a cognitive architecture, one wholly inconsistent with any form of intellectual intuition.[153] A reading of Gödel's work that took him somehow to 'blend' Kantian philosophy with a commitment to Platonic entities would owe a detailed account of how these clashing commitments could be reconciled. The endeavour is, I think, hopeless. Moreover, I find no evidence in Gödel's writings that he attempted such a reconciliation.[154]

Syntheses. At this point, I must ask for the reader's indulgence. Kurt Gödel's account of our access to the truths of mathematics turns on some of the details of the interplay between sensibility and intellect. This is not the place for a close reading of the *Critique* but it will not be possible to discuss Gödel competently without reviewing some of these.[155]

According to Kant, the raw presentations to the sensibility, whether these arise as the effects of mind-independent entities or from internal causes, contain what Kant calls a *manifold*. They are, in other words, complex and consist of a multiplicity of potentially distinguishable presentations. A problem lurks here. As we have seen, the passive reception of impressions as such is an immediate event that involves no computation. Yet in order for the received multiplicity of the manifold to potentially resolve into a diachronic unity in which distinct elements are distinguishable, each immediate presentation must be susceptible to being gathered as a single, unified presentation, however briefly. This is accomplished by what Kant calls a *synthesis of apprehension in intuition* [A99]. The details are a bit sketchy but what seems plausible is that the synthesis of apprehension relies passively on the spatio-temporal arrangement of the imprints to bring them into appropriate unity. (More on space and time in a moment.)

Next, the gathered manifold is subject to the simplest function of the spontaneous intellect [B130]. Kant reasons that in order for an experience of an object to be possible, the presented manifold must not only be gathered together but also become susceptible to being held fixed and ultimately compared with both itself and other experiences. The holding fast of the synthesized manifold is performed by a faculty that will play a major role in what follows: the *reproductive imagination*. Let me say at once that the label is exceptionally misleading (even for Kant). This faculty is not to be confused with the productive imagination—the more familiar faculty responsible for generating phantasms and daydreams. I emphasize the point because Gödel places a good deal of weight on this

---

occurs. This perhaps is the initial motivation for the transcendental deduction of the categories.

[153] See Falkenstein [1995] for an extended discussion.

[154] Martin Solomon [2009] proposes such a reading.

[155] I am grateful to Andy Brook for a number of conversations about the doctrine of synthesis. This section draws on his Kant seminar and his [1994] book.

particular process. For the moment, let us simply characterize the productive imagination functionally: it lies at the interface between brute sensibility and spontaneous intellect. It constitutes (on most readings) the lowest level of the spontaneity subject to the laws of association [A101,B151]. However, that said, it's somewhat unclear whether it manipulates sensations themselves or their representations.

The gathered and held manifold is subject to a third synthesis: *recognition in a concept*. This final act of synthesis subsumes the manifold under a general predicate resulting ultimately in a stable intentional object which is experienced as falling under a description [A104]. A good deal of work is done by the cognitive apparatus at this crucial stage. Importantly, the mere reproduction of the synthesized manifold in the imagination does not yet yield a fully fledged object of experience. In other words, there is more to being a recognizable, unified intentional object than being a bundle of representations upon which the repetition of the reproductive imagination is effected. At a minimum, it requires the recognition of the output of the reproductive imagination *as* a single object. The brute fact that the representations being synthesized into a unified intentional object (presumably) have as their source an objective, mind-independent, concrete particular does not help much. If the idea of an unified object itself were not antecedently available to the mind, at least in a purely formal sense, the impressions made on the sensibility would not themselves be sufficient to spontaneously generate it.[156] An act of judgement that connects the represented elements in a single experience is requited. (This is borne out by recent empirical work on visual apperceptive agnosia.[157]) Kant calls the act of judgement in which the mind connects disparate represented elements in a single experience the *unity of apperception.*[158] To effect it, the mind itself needs to be capable of applying a function that becomes aware of representations *as* a single experience. The apperceptive unity is effected on the basis of a transcendental object, a pure x (or object file, of you prefer) that is already in the mind. The application of the act of recognition in a single object itself requires the application of several kinds of concepts; at least those of quantity, quality, and modality. So there are built-in constraints on what intentional objects are like.[159] Finally, an unified intentional object falls under some more specific concept; it's recognized as a triangle, a teapot, or a face. The labels applied by the intellect are necessarily general; they do not relate us directly to concrete particulars but rather bring intentional objects under a general rule.[160]

[156]We will see Kant's point is reiterated by Gödel [1947].

[157]Shelton et al. [1994] discuss a subject with good visual acuity who nonetheless fails to recognize presented objects as unified wholes.

[158]See Brook [1994] for a helpful extended discussion.

[159]I will not address the transcendental deduction of the categories here since nothing in the two Gödel papers that we will shortly discuss draws on it.

[160]While the vocabulary is tortured and antiquated, the account is by no means out of date. Brook [1994] and Kitcher [1990]) both point out that Anne Treisman's influential [1980] model of feature binding exactly

The upshot of Kant's account of the interplay between sensibility and intellect is that, in the normal way of things, the two faculties are wholly codependent. The synthesis of recognition in a concept depends on the prior activities of the syntheses of apprehension and reproduction. And while the first two syntheses can be accomplished without the third, the product this yields cannot be an object of experience for us. All thought, indeed all phenomenal experience, is the result of the application of *both* faculties—which is to say of all three syntheses. The consistent application of this co-dependence thesis entails two weighty limitative results. First, the use of concepts divorced from application to content derived, mediately or immediately, from sensory experience yields nothing but a spurious spinning of the wheels. We can certainly manipulate ideas either by free-associating or by cogitating in accordance with the principle of non-contradiction. But, as Kant puts it, dialectical cognition bereft of sensible input must ultimately remain *empty.* Clearly, this limitative result constitutes a major challenge to anyone interested in defending entities or truths that are purportedly not originally intuited by the senses. In particular, metaphysicians and natural scientists who propose to violate the *emptiness thesis*—either by positing *ante rem* entities or even imperceptible atoms—owe an account of how minds such as ours can come by the relevant content. In the absence of an account, one must suspect that the hypothesized entities are merely phantasms generated by the productive imagination (a problem of which Gödel would have been aware).[161] The second limitative result cuts the other way: sensations that do not undergo the syntheses of the understanding and so are not subsumed under representations are *blind.* They are not devoid of content, but they do stand in a wholly non-cognitive relation to us and cannot directly become a matter of human thought or experience. Here we come full circle and see the justification for Kant's insistence that we have no privileged access to the nature our own constitution. Blindness prevents us from directly apprehending aspects of our being except by way of the recognition of sensory experiences in a representation.

Space & Time. Let me end the discussion of the Kantian cognitive architecture by presenting a doctrine that is important both to Kant's own view of the content of mathematics and also to Gödel's understanding of the topic; namely, Kant's conception of space and time. Patricia Kitcher [1990] argues that we risk misreading Kant here if we do not see him as engaging in the controversies of the day.[162] This is clearest concerning the doctrine of space in particular, where Kant weighs in on a (by then) old dispute. It was well recognized that *if* our sense of space were derived from the matter of our sensations, it would need to be attributed to either vision or to the sense of touch. Our olfactory and auditory faculties are

recapitulates the Kantian doctrine of synthesis.

[161]We have here the seeds of Ernst Mach's *phenomenalism* which both Einstein and Gödel rejected. See Yourgaru [2005].

[162]My reading here follows Kitcher's [1990] and is also informed by Falkenstein 1990, 1995.

not plausible as the origin of our experience of spatial relations. Neither sight nor touch however seemed suitable to the task. The retina is, in essence, a two dimensional surface and so cannot itself judge distance in three dimensions. Moreover, a variety of wholly different external objects can cause identical retinal presentations (even if we factor in binocular presentation). Descartes had proposed that our sense of distance could be accounted for by positing an 'innate geometry' which exploited the parallax due to the distance between the eyes. The proposal had been criticized, among others, by Berkeley.[163] And it ran into several significant objections: it was known, for instance, that the congenitally blind have a good sense of spatial relations.[164] Moreover, we typically experience dreams as presented to us in space. But, of course, visual parallax is of no use in such a context. Yet neither can our sense of space be attributed to the sense of touch, as Berkeley had maintained. In response to that hypothesis, Leibniz had pointed out cases of paralyzed individuals learning Euclidean geometry by sight alone. Counterexamples, it seems, exist to proposals on both sides of the debate. (The argument for the nonsensory origin is, if anything, clearer in the case of our experience of time.)

Kitcher reads Kant—plausibly, I think—to be following the *modus tollens* through to the valid conclusion: our experience of space must be attributed to a source other than mere sensory impressions.[165] But neither is it plausible that spatial relations can have their origin in the ideas and discursive functions of the intellect. Kant asks us to consider in connection with this hypothesis two congruent scalene triangles inscribed on a sphere, one in the north hemisphere and one in the south, such that they share a common base [Kant 1783 §13]. All of the intrinsic properties of these two triangles are identical; all the same predicates apply to both. Indeed, as far as the intellect alone is concerned, they are indiscernible. But of course they *are* distinct entities insofar as they cannot be made to overlap (except by rotating them in three dimensions). The example is even more telling once presented in three dimensions: a pair of three dimensional stereoisomers (such as a pair of gloves) cannot be brought to overlap at all. From this, Kant concludes that spatial differences are not merely intellectual differences. Spatial relations are real and so due to some aspect of reality given to the sensibility and not merely conceptual differences detectable by the intellect.

Having discounted both the matter of sensation and relations among ideas as the sources of our spatio-temporal experience, Kant offers his own suggestion. The truly striking aspect of our spatial and temporal experiences is their universal and necessary character. We can antecedently be assured that all of our outer experiences will involve objects located in three

[163] Berkeley's actual criticisms turn on a flawed view of the role of consciousness in cognitive processing, so I will not discuss them here.

[164] For modern evidence, see Landau et al. [1981] and Landau et al. [1984].

[165] Note that Kant may have reached an important and correct conclusion by means of a valid, yet unsound argument.

dimensional, Euclidean space. Similarly, we know that all of our inner experiences will be capable in principle of being ordered in a single, linear, temporal sequence (individual failures of memory notwithstanding). Indeed, we can imagine being deprived of all input to the sensibility but even under those circumstances, the spatiotemporal character of experience would persist. We have no such assurance concerning any other aspect of our sensations.[166] Kant infers that universal features of experience such as these must be due (not to the matter of sensation but) to the way in which sensations are themselves arrayed in the sensibility. Simply put, what I have been calling sensory experience is a complex of two elements: the matter of sensation (the sensa) and the order in the sensations are arrayed [Falkenstein 1990]. Space is the ordering principle of all of our outer experiences; time is the ordering principle of both outer and inner experience.

The solution is ingenious. It does however impose a further limitative result. Recall that the consistent application of the blindness and emptiness theses forces Kant to say next to nothing about objects as they are in themselves. And he is, if anything, even more circumspect about the nature of space and of time as these exist in themselves.

> I consider all the representations of the senses, together with their form, space and time, to be nothing but appearances, and space and time to be a mere form of the sensibility, which is not to be met with in objects out of it . . . But if I venture to go beyond all possible experience with my concepts of space and time, which I cannot refrain from doing if I proclaim them qualities inherent in things themselves. . . then a grave error may arise due to illusion, in which I proclaim to be universally valid what is merely a subjective condition of the intuition of things and certain only for the objects of . . . possible experience. [Kant 1783, P291-2]

Kant then does not permit himself to characterize space or time wholly apart from our experience of them. The ordering by means of which our sensations are arrayed as such has *no reality* whatever apart from human experience. (Gödel, we shall see, rejects the purely subjective character of space and time while accepting the essentials of Kant's account.)

Mathematics. It's not an exaggeration that explaining the striking contrast between the successful employment of pure reason in mathematics and its failure to yield useful results in metaphysics preoccupied Kant during some of the most active years of his life.[167] Regrettably, Kant was not fortunate enough to witness the period of rapid and profound mathematical development which occurred in the course of the nineteenth century [Kline 1980]. Even so, already in his time the difference in the fortunes of speculative philosophy

[166] The possible exception is that we know antecedently that all our experiences will be attributable to a single 'I'. I will not however address the transcendental unity of apperception here.

[167] By which I mean from the time of the *Inaugural Dissertation* of 1764 to the time of the 1783 *Prolegomena*. See Carson [2004].

and mathematics were plain to see.[168] The diagnosis Kant offers of the contrast between the disciplines is rooted his account of the faculties: in brief, the mathematicians' use of reason is different in kind from the philosophers'. Along the way, he offers an original and workable explanation of the nature of mathematical judgements.

The central phenomenon to be explained (by transcendental argument) is that mathematical research is genuinely ampliative. Kant was aware, of course, that mathematical progress is not always steady and that blind alleys do frequently arise. But the point is that mathematicians (like cartographers and experimental scientists) do build on prior results to push back the boundaries of our ignorance. Now, in general, ampliative judgements have this feature: they can contain genuinely novel information that cannot be gleaned by the reshuffling and re-association of ideas alone. This is not to say that analysis is unhelpful. On the contrary; careful scrutiny of ideas often uncovers hidden inconsistencies. And where our ideas can be analyzed and shown to violate the principle of noncontradiction we can be sure that something has gone wrong. Genuinely novel content however cannot arise from the discursive exercise of the intellect. In the case of empirical (*a posteriori*) judgements, new content comes from an encounter with the world. Indeed, the fact that the world forces itself upon us as being a certain way was, recall, among the motivations for inferring the existence of a receptive faculty in the first place. Similarly, to the extent that (some) mathematical judgements are ampliative (or *synthetic*) suggests that when making them we intuit content external to and independent of the intellect.[169]

Empirical ampliative judgements present no special explanatory problems for Kant: the receptive faculty receives impressions from without. Questions about the structure of that reality to which we lack answers can be pursued, straightforwardly enough, by creating conditions under which new experiences can be had. Thus, if we do not know whether a given substance conducts electrical current or whether it floats in ethanol, we create experimental conditions that afford us the appropriate intuitions to make the relevant judgement. In short, 'questions not decidable now may be decidable in light of additional experience.' Recall however that according to Kant, all intuition is sensible intuition; that is, all new content is derived from sensory experience. It seems wholly implausible that the content of mathematical judgements derives from the content of our sensations. One reason Kant offers against such an explanation is that mathematical judgements are *universal.* When we

[168]In fact, this is a lesson Kant learns from Hume [1748]. It is a lesson however that has apparently not been fully absorbed by some modern philosophers.

[169]Kant claims that all mathematical judgements are synthetic [A11]. The point is debatable. I think a more careful way of putting it is that all mathematical judgements are synthetic *in their origins.* That is, all new mathematical content, when first derived, is genuinely synthetic or ampliative. However, mathematical results can be justified analytically by means of formal proof from given assumptions. Kant's writing historically antedates the heavy emphasis on formal proof that became important with the foundational work of the nineteenth century so I think we ought to read him charitably here. The important point is the tight link between truly *new* mathematical content and synthetic judgement.

learn that a triangle inscribed in a circle (such that the base of the triangle bisects the circle) contains a right angle, we learn this *for all* inscribed triangles. There is no need to conduct additional experiments. Unlike empirical judgements, mathematical judgements do not depend on any *particular* experience for their justification. Secondly, correct mathematical judgements are *necessary*: they cannot be altered and are a feature of any experience.[170] How then are pure (by which I mean non-empirical), ampliative judgements justified *a priori*?

Given what has been said, the answer is obvious: there is more to sensory experience than sensations alone.[171] As we just saw, in addition to the content of the sensations, sensory experience embodies also the principles of organization. Sensations are arrayed in the intuition both spatially and temporally. And it's precisely this ordering of the matter of sensible intuitions which, Kant suggests, itself becomes the object of mathematical reflection. In the case of arithmetic, the content derives from the pure, *a priori* intuition of time [A720]. Geometric judgements on the other hand are rooted in our *a priori* intuition of space. When reflecting on geometric facts, the geometer effectively constructs a figure

> by representing the object which corresponds to this concept either by [reproductive!] imagination alone, in pure intuition, or in accordance therewith also on paper, in empirical intuition—in both cases completely apriori, without having borrowed the pattern from [the matter of] any experience.[A713]

In the case of mathematical cognition, our ideas arise from individual experiences but give us universal knowledge about any experience whatsoever, provided that experience is governed by the same principles of organization. (Let me make one note here: this passage is sometimes read as referring to the productive imagination. This cannot be right. There's no barrier to the productive imagination forming phantasms that, while apparently convincing, are incoherent: one can perfectly well picture by means of the productive imagination using a ruler and compass to nudge a unit circle into a perfect square. Yet, of course, squaring the circle is impossible. Kant was not ignorant of this. What's at issue here is not our capacity to form fanciful phantasms. Rather, it's our capacity to synthesize factually occurring spatial and temporal patterns that does the cognitive work. This is done by the reproductive imagination [A102]. Its products may sometimes be mirrored in phantasms, but need not be.) The necessity of mathematical judgements derives from the perfect generality of the forms of the sensibility. All content of sensory intuition is arrayed in space and time in just the same way. If mathematics derives its basic content by (unselfconsciously) studying that structure, we can be sure that the most basic mathematical results will apply universally to

[170] This is not to deny that we sometimes make mistakes in recognizing which mathematical judgements are true.

[171] Falkenstein's [1990, 1995] work provides a particularly lucid and helpful discussion of this.

sensory experience. It follows unfortunately that mathematics will not legitimately apply to entities as they exist in themselves, apart from our experience of them. If science relies essentially on mathematics then, for Kant at least, a science of the noumena is impossible. (We will see that Gödel takes a different view.) The possibility of mathematical error is explained in just the same way that perceptual error is accounted for: our empirical judgements are sometimes false because, on occasion, the ideas under which we subsume presentations do not match the ideas we would have subsumed presentations under had we had more exposure to the manifold in question.[172] Mathematical definitions are attempts to subsume under concepts the universally valid content present in the reproductive imagination.[173] And sometimes these attempts eventually prove incoherent.

Metaphysics. Mathematical judgements can be explained entirely a priori, with no reference to the matter presented to the sensibility. It may still seem puzzling though that the employment of pure reason in mathematics is cumulative, while its employment in metaphysical research is not. Let's review the reason why.

The mathematician proceeds from intuitions to concepts, constructing the latter on the basis of the former (or at least trying to do so). As Kant puts it, she "contemplates the universal in the particular" [A714] by discovering universally valid patterns in the spatio-temporal array by way of which the matter of sensation is presented. Though Kant does not say so, one might add that as research proceeds, the construction of ever more complex such patterns (of spatio-temporal configurations) becomes possible—a sort of evolving, increasingly intricate kaleidoscope on which to base new ideas. The mathematical realm is therefore real, but essentially unfinished. The speculative metaphysicians' employment of pure reason is altogether different [A713-738]. Speculative metaphysics, Kant suggests, does not construct concepts on the basis of new intuitions. Instead, it "contemplates the particular in the universal." That is, it proceeds by taking a particular entity and analyzing the concepts that apply to it. This is a discursive activity of the intellect and so can generate no new content. In fact, contemplating an individual object—say, a triangle—and enumerating the concepts under which it falls cannot help us prove so much as a single theorem of Euclidean geometry. If the reception of content from outside the intellect is the engine of conceptual progress and if the mathematician is able to do this while the metaphysician cannot, speculative metaphysics must remain nothing more than a rehashing of ideas.[174]

[172]A special case of this are optical illusions and bistable percepts; in such cases, the same intuition can equally well be subsumed under two distinct concepts. It's at least conceivable that illusions of a sort arise in mathematics. Gödel (below) takes this possibility seriously.

[173]In some passages [A713-738], Kant takes a different view of definitions. Given what he says about mathematical content and its independence of the intellect, I think this is a mistake. In any case, Kant's views have been made obsolete by subsequent dicussions—notably between Frege and Hilbert—concerning definitions. For a useful discussion, see Brown [1999].

[174]I have found Emily Carson's [1999] work very helpful in thinking about this section.

Gödel's Transcendental Realism

Godel's explanation of our knowledge of mathematics recapitulates Kant in many respects. Like Kant's explanation, it ultimately rests on Gödel's conception of space, time, and the human cognitive architecture. Remarkably, Gödel maintains that Kant's doctrines on these topics are largely correct and in accord with modern science—though they need to be interpreted carefully and supplemented to reflect recent work. The reading of Kant that Gödel pursues places a great deal of weight on two aspects of the philosophy: the unequivocal rejection of subjective idealism, and the distinction between the manifest world of conscious experience and reality as it exists in itself. If we hang on firmly to these two insights (and allow ourselves to reassess flexibly some of the others) almost the whole of Kant's metaphysics remains viable. The unusual reading of transcendental idealism which Gödel offers may not please a scrupulous scholar or an orthodox Kantian; it does however give a slightly more opportunistic natural philosopher a way of extending NORM in a genuinely new and untried direction.[175]

Science & Noumena. Let me begin with what Gödel sees as the single major point of *dis*agreement between the modern physics and Kant. It's useful to discuss it early since it will help to situate us. The three limitative results we have encountered—blindness, emptiness and the transcendental ideality of space and time—are typically taken to preclude any robust theory of entities as they exist in themselves. On most readings, Kant himself thought so.

> I say that things as objects of our senses existing outside us are given, but we know nothing of what they may be in themselves, knowing only their appearances; i.e., the representations which they cause in us by affecting our senses. Consequently, I grant by all means that there are bodies without us, that is, things which, though quite unknown to us as to what they are in themselves, we yet know by the representations which their influence on our sensibility procures us, and which we call bodies.[Kant 1783, Part 1, Remark III.]

An important consequence of the limitative results, according to Kant, is that scientific disciplines—including physiology, cosmology, and chemistry—may legitimately inform us about the range of possible experiences we may undergo, but never about mind-independent reality. This doctrine, *phenomenalism*, has an unfortunate past. Partly due to Kant's influence, it became a powerful intellectual force during the nineteenth century, culminating in the Ernst Mach's steadfast and effective opposition to atomic theory. Gödel flatly rejects

[175] As it happens, Gödel himself was not a naturalist. Neither, for that matter, was Kant. But one of the great benefits of enlightened Kantian philosophy is that theological issues can be left as a matter for personal conscience. That is to say, the account of mathematical content offered here is consistent with ontological naturalism as well as with deism (and perhaps even some varieties of theism).

this view.[176] It may well have been the case, he argues, that physics in the time of Newton was limited to discovering and clarifying "relations between appearances." However, this has now changed.

> The abandoning of that 'natural' picture of the world which Kant calls the wold of 'appearance' is exactly the main characteristic distinguishing modern physics from Newtonian physics. Newtonian physics, except for the elimination of secondary qualities... is only a refinement, but not a correction, of this picture of the world; modern physics, however, has an entirely different character.[C1,27]

According to Gödel, Newtonian physics was, in principle at least, translatable into a regimented laboratory language (perhaps in the way envisaged by Carnap). Today, Gödel argues, we are not so limited. What permits modern scientific work to move beyond refined descriptions of phenomena is precisely an increased reliance on a technical vocabulary that is ultimately cashed out in purely mathematical terms. By degrees, and in a slow and groping manner, Gödel suggests, our best theories thus afford us a glimpse of what things are like in themselves [C1,27].

Time & Space. Gödel does not see his views on noumenal access as a truly radical break with Kant. Rather, he sees his interpretation as drawing out some of the hidden potential of Kantian critical philosophy, modifying inessential aspects, and remaining true to its spirit. On Gödel's reading, for instance, Kant was himself intermittently open to the possibility of a knowledge of noumena, though "he wanted to base such knowledge on ethics."[177] Indeed, Gödel claims to find in Kant's discussions of subjective idealism brief moments when Kant permits himself to comment on the nature of noumenal reality. Here is the main such passage Gödel discusses:

> Suppose ... that I could intuit myself without being subject to this condition of sensibility [i.e., time] or that another being could so intuit me; in that case the very same *modifications that we now represent to ourselves as changes* would provide a cognition in which the presentation of time, and hence also that of change, would not occur at all. [Kant 1781, B54, translation modified to match Gödel's]

On Gödel's reading, the 'modifications' of which Kant speaks here cannot be understood as features of our conscious lives. For the passage to make sense, Kant must be referring to objective analogues of the phenomenal temporal changes that we experience. In Gödel's view, the passage is not a momentary lapse of attention on Kant's part. Indeed, Kant cannot

[176] As, of course, does Einstein whose work on Brownian motion contributed significantly to the acceptance of atomic theory and hastened the demise of Mach's conception of physics. For some of the background and a useful discussion see Yourgrau [2005]. For a modern throwback to those times, see van Fraassen [1998].

[177] In support of his contention, Gödel cites the preface to the B edition of the *Critique*: B xxi, B xxii n. and B xxvi n. See [Gödel 1949, C1,29-30].

concede that space and time are wholly subjective illusions if he is to escape the charge of subjective idealism. Insisting merely that space and time are ubiquitous features of experience would not help; as a persistent delusion is nonetheless a delusion. On Gödel's reading Kant therefore quite deliberately intends to commit himself to the view that the spatial and temporal properties of entities that we experience have an objective, mind-independent analogue. Kant is, of course, pessimistic about being able to accurately characterize that mind-indepedent analogue and so does not try to do so. Naming the enigmatic 'modifications' mentioned in the above passage is as far as his commitment to the limitative theses allows him to go.

Taking the discussion of temporal 'modifications' as his starting point allows Gödel to argue that nearly everything about Kant's conception of time is consistent with general relativity—indeed, that the former is a sort of inchoate forerunner of the latter. This may seem like a hard claim to accept. Gödel is suggesting, after all, that passages such as the following, are wholly in accord with his own realism *and* that they remain scientifically respectable today:

> Time is nothing but the form of our inner intuition. If we take away from time the qualification that it is the special condition of our sensibility, then the concept of time vanishes as well; time attaches not to objects themselves, but merely to the subject intuiting them.[Kant 1781, B54]

The paradox, on Gödel's reading, is apparent rather than real. It's easily dispelled once we unambiguously distinguish (as Kant should have done) between two meanings of 'time'. This, Gödel suggests with characteristic understatement, leads to a theory that is "slightly different" from the one offered by Kant, but nonetheless consistent with the overall thrust of his philosophical vision.[178] In order to keep the terminology straight, let me here explicitly distinguish between time$_\phi$ and time$_\nu$. Time$_\phi$ is time as creatures such as ourselves experience it: phenomenal time.[179] Time$_\nu$ by contrast refers to those modifications of the things in themselves that correspond to our sense of temporal change; noumenal time, if you like. According to Gödel's interpretation, in nearly all passages in which Kant discusses time, he should be read as talking about time$_\phi$—which is to say, the temporal structure of conscious experience. On this reading, the above passage turns out to be perfectly correct; indeed, it's nearly a truism. There plausibly is nothing more to time$_\phi$ than the form of our inner intuition. And if we were to abstract the peculiarities of our receptive faculty, time$_\phi$, would be washed out as well. Finally, it would be a mistake to attribute many of the features of time$_\phi$ to modifications of objects existing in themselves, just as Kant had maintained.

Gödel takes up the contrast between the features of time$_\phi$ and time$_\nu$ in detail. Time,

[178] An interesting precursor of this idea can be found in Weyl [1927].

[179] I assume without argument that many nonhuman animals experience time$_\phi$ in just the way we do.

understood now as the consciously experienced realium, is experienced as "flowing." Admittedly, the river metaphor is a bad one but it's notoriously difficult to do better. Time$_\phi$ *seems* to flow at variable pace depending, in part, on our physical state, level of alertness and the activities we are engaged in, though on reflection we know that this variability is really a sort of illusion; really, time$_\phi$ as such moves at a uniform pace throughout the world. These days, educated common sense construes time$_\phi$ as having a linear structure and as being unbounded both into the past and into the future.[180] It's possible to characterize time$_\phi$ more formally as well though. Construed as a relation between events, time$_\phi$ constitutes a *total*, *transitive*, *anti-symmetric* order.[181] In other words, all events whatsoever can be placed in a temporal sequence relative one another; if A occurs before B and B occurs before C then it follows that A precedes C; and, no two distinct events can precede each other. One interesting consequence of this view, Gödel notes, is that at any instant in time, only a subset of the facts that, considered as a totality constitute the *cosmos*, actually exists. Empirical entities are therefore experienced as coming to be and passing away in time$_\phi$.[182]

The noumenal modifications "which we represent to ourselves as changes" constitute a mind-independent correlate of our subjective sense of time. Time$_\nu$ cannot be experienced directly. But, Gödel suggests, it has successfully been characterized by modern physics, and by the Lorentz transformation of relativity theory in particular [Einstein 1961]. Unlike its phenomenal counterpart, time$_\nu$ does not flow and is not linear. To characterize the temporal relationship between two events, we need to specify a frame of reference (so also their velocities and accelerations relative that frame of reference). Specifying the temporal coordinate of an event thus requires (not one but) four coordinates. Temporally related events stand in no fixed relation: given two events A and B, which precedes the other depends on the reference frame of the observer. It's not possible therefore to localize the event in time$_\nu$ using a simple number line. Moreover, time$_\nu$ does not constitute a total ordering of occurrences since, in a relativistic universe, some events may stand in no temporal relation to one another. As I noted earlier, Gödel was able to show that relativity permits the existence of worlds in which no simultaneity relation can be defined. Lastly, unlike the time of common sense, time$_\nu$ it is bounded in the past and perhaps also in the future by cataclysmic cosmic events. On the whole then, time$_\nu$ is quite unlike its phenomenal counterpart.

Gödel's treatment of space is broadly similar to his treatment of time.[183] Consider the

---

[180] This has not always been a part of educated common sense. Pylyshyn [1984] traces our modern conception of time$_\phi$ to Galileo, though Kant would surely disagree. I leave this as an open question.

[181] One could add also that common-sense time is construed as *acyclical*, thereby excluding the Stoic conception of the temporal order. For an interesting recent discussion of the Stoics see Sakezles [2009].

[182] What Kant calls 'pure' entities, including those of mathematics, are atemporal$_\phi$. It does not follow however that they are atemporal in some stronger, mystical sense, as the *ante rem* realist holds.

[183] I am glossing over some inessential subtleties here. The important point is that once Gödel recognizes that quantum theory posits a space$_\nu$ wholly unlike our space$_\phi$ [B2,13] whatever differences there may be

following passage from Kant:

> Space represents no property whatever of any things in themselves, nor does it represent things in themselves in their relation to one another. That is, space represents no determination of such things, no determination that adheres to objects themselves and that would remain even if we abstracted from all subjective conditions of intuition. [A26]

Introducing a distinction between a space$_\phi$ and space$_\nu$ is once again helpful here. With the distinction in place, the passage once again reads like a near truism. And, once again, the properties of space$_\phi$ and space$_\nu$ contrast sharply. All space we are capable of experiencing is three dimensional and Euclidean. It is infinite in every direction. Moreover, space$_\phi$ is unified and homogenous, containing no gaps or lacunae.[184] Finally, space$_\phi$ is infinitely divisible both conceptually and (we imagine) physically as well. Famously, space$_\nu$ is not like that. General relativity teaches that space$_\nu$ is non-Euclidean and that its curvature varies with the local mass. The spatial$_\nu$ dimensions of objects depend on the frame of reference from which we view them, though this is only evident when extreme velocities are involved. Moreover, space$_\nu$ may well prove unbounded but not infinite. And finally, space$_\nu$ there may be a lower bound on the divisibility of physical space. Here once again then we have a contrast between features of reality as we experience it and what our best physics tells us concerning reality [Einstein 1961].

Adopting a realistic conception of the noumena—including space$_\nu$ and time$_\nu$—and a robustly optimistic view of our capacity to come to know them sets Gödel apart from orthodox Kantians. Gödel also thereby incurs some nontrivial debts. If he is to remain consistent, he must either reject the Kantian conception of the mind—the source of the limitative theses—or he must show that his robust realism can be reconciled with blindness and emptiness.[185] Just as importantly, he needs to explain how it's possible for mathematics—which, of course, Kant had grounded in the *subjective* forms of the sensibility—to help us peek beneath the veil of phenomena. By clarifying how Gödel meets this challenge we come *ipso facto* to understand his treatment of human access to mathematical facts.

Dual access. Though this is tendentious, I think Gödel [1949] is aware of his philosophical obligations. Some of the evidence is circumstantial: I've already noted the depth of Gödel's

between the treatment of space and time become inessential.

[184]The unity and continuity of space$_\phi$ can be contested; the space of our dreams is not continuous with the space$_\phi$ experienced while awake. This however only lends support to the notion that neither vision nor touch are necessary for spatial experience. See above.

[185]As tempting as rejecting the Kantian cognitive architecture might be, it requires one to offer an alternative explanation of the applicability of mathematics to the phenomena of experience, as well as to explain from scratch our access to mathematical truths. A good deal of modern philosophy of mathematics can be read as the rejection of Kantian transcendental psychology and the subsequent struggle to address those two problems without the benefit of Kantian insights.

understanding of Kantian philosophy. I've also noted how assiduously Gödel worked on the [1949] paper. Ultimately however, the proof must be in the text. As I read him, Gödel deliberately makes several subtle yet significant adaptations to the classic Kantian cognitive architecture so as to overcome the limitative theses.[186]

The principal such modification concerns the interpretation of the reproductive imagination's place within the cognitive architecture. Several commentators have noted that this sub-faculty plays an ambiguous role in Kant's overall scheme; it's unclear whether it's to be understood as the most sophisticated aspect of a passive receptive faculty or the most rudimentary aspect of the spontaneous intellect.[187] In other words, it's not clear whether we should construe it as passively reproducing transduced impressions or actively computing over basic, structured representations. Earlier, I defaulted to the latter reading. Gödel opts for the former. That is, he opts to include the reproductive imagination within the purview of the sensibility rather than construing it as part of the intellect. The relevant passage occurs in Gödel's discussion of the crucial [B54] discussion concerning the existence of objective temporal$_\nu$ 'modifications':

> in this passage a view as to the nature of space and time, slightly different from that usually ascribed to Kant, seems to be implied, which however is not incompatible with the latter insofar as, corresponding to the *two parts of the sensibility* (the faculty of sensation and of representation), both kinds of relations of the things to sensibility may subsist beside each other.[B2,4b, added emphasis]

This may read like a cosmetic change or a trivial detail. In fact, it makes an important difference. Recall how the *emptiness* restriction is characterized: intellectual concepts are empty, Kant holds, if they are generated by the rehashing of ideas in the intellect and float unmoored from the operations of the sensibility. If both the reproductive imagination and the passive receptivity can be construed as semi-autonomous aspects of the sensibility then, plausibly, each can play an independent and distinct grounding role. Indeed, they can "subsist beside each other" and act independently to anchor experience. The bifurcation of the sensibility helps explain some of Gödel's otherwise puzzling commitments. According to him, we stand in two relations to noumenal entities, their properties$_\nu$, and relations$_\nu$: the one is (what Gödel calls) cognitive; the other is brute and merely factual. We likewise have two modes of access to geometric facts. We know the truths of Euclidean geometry to hold a priori; at the same time, "geometry is in one sense an empirical science" [4b]. Let me try to disentangle these views.

---

[186]Not everyone agrees. In his helpful commentary on Gödel [1949], Howard Stein, for instance, places a great deal of weight on Gödel's assertion that he is not a Kantian and consequently does not read Gödel as attempting to meet blindness and emptiness constraints.

[187]See Brook [1994] for a discussion.

Let's start with our cognitive access to space$_\phi$ (and time$_\phi$) and our a priori knowledge of Euclidean geometry. Gödel agrees with Kant that, in practice, the forms of our sensibility guarantee that human experience will be as of a three dimensional, Euclidean space (though see below). Any spatial relations we can expect to encounter are sure to be like that since we are "able to form images of outer objects only by projecting our sensations on this representation of space" [B2,14]. We can thus be certain a priori that Euclidean geometry will apply to space$_\phi$ as we experience it. I take Gödel here not to be introducing any changes to the Kantian account. And, according to Kant, the reproductive imagination plays an integral role in the phenomenal projection of even the simplest experiences, including that of space$_\phi$. Without the synthesis of reproduction, "there could never arise a whole presentation: nor could there arise... even the purest and most basic presentation of space and time" [A102]. On this reading, intellectual representations of Euclidean geometric figures are anchored in the sensibility in the standard way. A mathematician or physicist contemplating (say) a triangle relies on the forms of the passive receptivity, the synthesis of reproduction in the imagination, and finally a recognition of the presentation under a concept. Note that this is consistent with general relativity's curved space-time$_\nu$ since Euclidean relations can—Gödel points out—be defined in a non-Euclidean space, though not as an absolute but rather in relation to a privileged coordinate system. If we take the sensibility (construed now as sensory organs arrayed in space) as that privileged object, we see why Gödel took Kant to have been a forerunner of Einstein. Gödel writes:

> for Kant [space] is a relation to his 'sensibility' (which presumably means that it depends on the special structure of his organs of sense and representation) [B2,8]
>
> our representation of space is completely adequate to the relation which our sensibility has to the objects.[B2,13].

To the orthodox Kantian, Gödel's view offers rather hollow vindication. For while we have a priori access to the truths of Euclidean geometry as well as assureance that this geometry will correctly describe objects as we expereince them, we lose any guarantee that such descriptions offer us any insight into what the world is like *per se*. Indeed, it remains an open question, Gödel suggests, whether our spatial$_\phi$ experience would be similar in a strongly non-Euclidean world. It's possible, he suggests, that under such conditions we would nonetheless 'imagine,' or rather project, space$_\phi$ as homogenous and Euclidean. He speculates that what would change under those circumstances is our conception of the motion of rigid bodies. Plausibly, we would experience solid objects changing their size and shape as they moved through an apparently uniform, Euclidean medium [B2,14].

How then do we come to access spatial$_\nu$ relations, and in what sense is geometry a posteriori? I must admit that my reconstruction here becomes somewhat speculative. I take it however that the essence of the proposal is a direct grounding of some mathematical content in the activities and potential configurations$_\nu$ inherent in the reproductive imagination. Here's one way this might work: if the reproductive imagination is construed as an aspect of the sensibility distinct from the apprehension then the patterns$_\nu$ of activity of the reproductive imagination can themselves become subject to a (secondary) synthesis of apprehension; that is, they can themselves be taken up as the manifold that is held fast as a unity. This is possible, note, only insofar as the reproductive imagination is construed—following Gödel—as a brute, factual process, and therefore as capable of serving as an independent ground for contentul cognition. Construed thus, the reproductive imagination is (according to a transcendental realist) a noumenal process like any other, with all the complexity and spatio-temporal$_\nu$ structure this entails. Once such a secondary synthesis of apprehension is accomplished, the subsequent cognitive processes proceed as they would in the case of the cognizing of a external, empirical object. The held manifold is itself subject to the synthesis of reproduction in the imagination (perhaps straining existing resources). And the result is classed under a rule or concept—though with a difference: in the Transcendental Doctrine of Method, Kant argues that empirical concepts cannot be defined but only spelled out [A728,B756]. That's because we "can never be sure that the distinct presentation given to [us]. . . has been developted comprehensively." The content derived via the second path at issue here must be regarded as empirical in just this sense. At best, our concepts of the content involved are expositions that hope to capture the underlying reality; they are not definitions that settle the matter by fiat. Our cognition of (say) non-Euclidean, higher dimensional geometry is—if not wholly *blind*, in the Kantian sense—then severely myopic. Research has the character of a groping the dark by means of imperfect concepts and without the benefit of a rich phenomenology provided by the experience of external things.[188] On this reading, geometry turns out to be "an empirical science" in two senses. On the one hand, *which* geometry turns out to apply to space$_\nu$ is an empirical issue. Just as importantly, it's not always initially clear which posits truly exist and which are mere phantasms.[189]

Once the details are laid out it becomes plain that neither of the Kantian limitative theses is violated in either of the two cases. And so Gödel is able to discharge his philosophical debts. We are offered the rudiments of Gödel's original, quasi-Kantian explanation of human mathematical knowledge. And we are offered an account of why mathematics furnishes us with (unsteady and hard-won) noumenal access. The successful conduct of this science depends on the construction of ideas of the intellect which accurately correspond to the factual states of the passive sensibility (itself a noumenal object). This does not violate emptiness since both the intellect and the sensibility play their respective roles. It does not

[188] The lack of phenomenology in such cases is an interesting problem, one which I leave unaddressed here.

[189] Recall in connection with this the dispute between the constructivist and the classicist concerning completed infinities discussed in Chapter 1. Recall also that Gödel advocated accepting or rejecting set-theoretic axioms depending on their future success. This is of a piece with the reading offered here.

violate blindness since, as a matter of fact, our non-Euclidean, geometric notions *are* nearly blind. Geometry is ampliative in either case; its content derives from intuition rather than by means of discursive cogitation in accordance with the principle of contradiction. It's precisely by means of mathematics (including non-Euclidean geometry) that we gradually and laboriously develop a sense of the world as it is in itself.

Set-theoretic intuition. Of course, it was not geometry but rather set theory that Gödel took to be the foundation of mathematics. And so, explaining our access to the facts of mathematics ultimately required him to offer an account of our knowledge of set theory. This is what Gödel sets out to provide in his celebrated [1947] essay. Before moving on, let me highlight just how closely Gödel's explanation of our access to the universe of sets parallels his treatment of our access to geometry. As we are about to see, it turns out to be harder to explain how the notion of membership can be explicated in terms of spatial or temporal relations. Yet even if we ultimately judge Gödel's attempt to explain our access to the primitives of set theory to be a failure, this may not matter. Many mathematicians and philosophers today, including Mac Lane [1997] and Hellman [2001], are willing to seriously entertain the possibility of alternative foundations for mathematics. It remains an open question therefore whether the Kantian account of epistemic access to fundamental mathematical content remains viable.

Gödel begins the main argument in the [1947] paper by offering a (valid) argument for the existence of a faculty of mathematical intuition. Here's how that argument runs. Let us say, following Kant, that if a judgement embodies genuinely novel content then it is ampliative (or synthetic, if you prefer). It can easily be shown that touch and visual perception can give rise to ampliative judgements. Recall why: there exist unsettled questions concerning visual and tactile properties. Limited as we are, we cannot settle these by the application of willpower alone. The contents of our visual or tactile perceptions 'force themselves upon us as being true.' Nor can undecidable perceptual questions be settled by thinking carefully about the issues. This is because the spontaneous intellect, left to its own devices, is capable of analytic judgements alone; it can recombine ideas and test them for mutual coherence, but it "cannot create any qualitatively new elements."[190] Contrary to subjective idealism, the intellect cannot therefore be the source of any new experiential content. And so, if giving answers to open perceptual questions is to be possible at all, it must be due to the gathering of additional experiential evidence. And since we settle such questions all the time, we can conclude that vision and touch are genuinely ampliative. What makes the argument interesting is that a structurally identical version can be used to show that mathematics generally, and set theory in particular, is also ampliative. The basic content of set theory, which is to say its principal axioms, "force themselves upon us" as being one way rather

[190]Note that this absolutely central Kantian insight squares with the a computational theory of mind. It also hints at its limitations.

than another. Alternative axiomatizations are clearly possible, but only in the sense that alternative perspectives on the same landscape are possible; dreaming up *arbitrarily different* axioms is beyond our power. Moreover, as with sensory experiences, there are genuine surprises in mathematics; questions that may have no answer at one moment in history (such as the independence of CH) can be determined in light of subsequent experience. If the source of our ampliative judgements cannot be attributed to analysis and recombination of ideas in the intellect then such judgements must involve input from without. Any such input requires the participation of appropriate receptive faculties—the faculties that, as we saw earlier, Kant neutrally labels *intuitus*. It follows therefore that we possess some form of mathematical intuition (in the Kantian sense).

Here are all the pieces of that same argument in Gödel's own words:

> Despite their remoteness from sense experience, we do have something like a perception of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e., in mathematical intuition, than in sense perception, which induces us to build up physical theories and to expect that future sense perceptions will agree with them and, moreover, to believe that a question not decidable now has meaning and may be decided in the future . . . [B]y our thinking we cannot create any qualitatively new elements, but only reproduce and combine those that are given. . . [And so] the set-theoretical paradoxes are hardly any more troublesome for mathematicians than deceptions of the senses for physics...[Gödel 1947]

Having satisfactorily established (in accord with Kant and contrary to the positivists) that mathematics is ampliative rather than analytic, and hence that it is based on some form of contact with outside reality, Gödel moves to consider the nature of that contact. The most obvious possibility for Gödel to pursue is that our set theoretic knowledge is simply a species of empirical knowledge; in other words, that it derives directly from the content of sensations themselves. Gödel does not seriously contemplate this proposal. The reason, I think, is that he considers many pure mathematical judgements to be a priori. If that's correct, then the content of set theoretic judgements cannot be traced to the content of sensations alone. Anyway, there's a better account available. As we saw earlier, on the Kantian view, our experience of objects generally—including mundane medium-sized, dry goods—derives in part from what is presented in intuition and in part from subsequent mental processing. In particular, the recognition of synchronically and diachronically unified empirical objects depends on the three syntheses as well as on the unity of apperception. The syntheses however depend on the forms of the sensibility. Thus even the simplest sensory experiences contain elements due both to the presentation of content in intuition (i.e., the sensations) and to the arraying of impressions in space$_\nu$ and time$_\nu$. If not for the latter, the synthesis of apprehension and of reproduction would be impossible. Here's Gödel discussing this issue:

> That something besides the sensations actually is immediately given follows (independently of mathematics) from the fact that even our ideas referring to physical objects contain constituents qualitatively different from sensations or mere combinations of sensations, e.g., the idea of the [spatio-temporal] object itself, whereas, on the other hand, by our thinking we cannot create any qualitatively new elements, but only reproduce and combine those that are given.[Gödel 1947]

Gödel wants to stress that while the raw material for mathematical judgements is intuitive, this does not mean that the sensibility *as such* places us in unmediated contact with sets in the same way that it places us in contact with concrete particulars. Gödel is perfectly comfortable making claims about noumenal reality. But that reality according to him, as we have seen, is the one revealed by modern physics. The iterative hierarchy is not among the posits of the general theory of relativity or any other branch of physics. So Gödel explicitly rejects the notion that our mathematical 'perception' should be thought of as offering us unmediated impressions (*Eindrucke*) of sets (whatever that would mean). Instead, just as in the case of phenomenal experience more generally, the sensibility is presented with content which, after the appropriate acts of synthesis, yields an experience of the objects$_\phi$ of mathematics. "[A]s in the case of physical experience, we form our ideas also of those objects on the basis of something else which is immediately given." The important thing to remember, both with respect to the objects encountered in phenomenal experience and the landscape met in mathematical investigation, is that while their properties may not be noumenally real, they are nonetheless adequate to the relation that we bear to their sources. The phenomeal experience we have—whether that's of empirical objects or of mathematical ones—is non-arbitrary and highly constrained. Here's Gödel again:

> Evidently the "given" underlying mathematics is closely related to the abstract elements contained in our empirical ideas... It should be noted that mathematical intuition need not be conceived as a faculty of giving an immediate [read: unmediated] knowledge of the objects concerned. Rather, it seems that, as in the case of physical experience, we form our ideas also of those objects on the basis of something else which is immediately given. Only this something else here [in the case of mathematicalia] is not, or not primarily, the sensations. [Gödel 1947]

It might be objected that we have too little evidence to identify the "abstract elements contained in our empirical ideas" with the forms of the sensibility. I think, given our earlier discussion of geometry, this is a weak objection. In any case, I believe my analysis is confirmed by the criticism that Gödel offers of Kant at the close of his discussion. Here is that passage:

> Evidently the "given" underlying mathematics is closely related to the abstract elements contained in our empirical ideas. It by no means follows, however, that the data of this second kind, because they cannot be associated with actions of certain things upon our

> sense organs, are something purely subjective, as Kant asserted. Rather they, too, may represent an aspect of objective reality, but, as opposed to sensations, their presence in us may be due to another kind of relationship between ourselves and reality. [Gödel 1947]

"The data of the second kind" cannot be identified with impressions transduced at the sensory surfaces. Nonetheless, the data is not subjective, Gödel insists, contrary to what Kant asserted. This, I think, cannot be read as anything other than a reaffirmation of Gödel's scientific realism concerning space$_\nu$ and time$_\nu$ and a swipe at Kant for limiting himself timidly to space$_\phi$ and time$_\phi$. As we have now seen, the justification for Gödel's view here can be found in his work on general relativity; he does not bother to review it here.

The reading works: Gödel, as we saw earlier, thought that our sensibility, construed as a physical entity, stood in a factual (non cognitive) relationship to other noumenal objects in space$_\nu$ and in time$_\nu$. It's the representation of this relation that, on my reading, he uses to ground our mathematical knowledge. Far from being the *cri de coeur* of a befuddled mystic then, the Gödel [1947] essay embodies a sober modern reinterpretation of Kant's solution to the problem posted by mathematical judgements in light of developments in modern physics and in the foundations of mathematics. At its core lies a responsible and serious realist proposal, albeit one that has been misunderstood and so largely sidelined.[191]

Knowledge of architecture

Our initial motivation, recall, was to develop a theory of structure (or grammar) for the human mathematics faculty. Gödel's philosophy of mathematics does not give us a ready-made solution to that problem. It does however give us a sense of how to understand the nature of strongly equivalent models of human mathematical cognition. It also sheds some borrowed light on the philosophical desiderata discussed earlier on. Before wrapping up the chapter, let's return to those.

Functional architecture. It will be helpful to translate Gödel's somewhat archaic vocabulary into a more modern, cognitive idiom. The Kantian *reproductive imagination* is a rather puzzling posit. Let's set aside the label for the moment and consider this sub-faculty from a strictly functional perspective. As we saw, the reproductive imagination is an element of the cognitive architecture that can legitimately be described both as an aspect of the receptivity *and* of the intellect. It thus occupies a unique position in the cognitive economy: in one sense, it is a brute physical process; yet it can be viewed simultaneously as manipulating (very basic) syntactically-individuated data-structures. Modern cognitive science makes use of a functionally nearly identical posit to ground analyses. Recall how explanation in cognitive science traditionally proceeds [Pylyshyn 1984]. A cognitive ability

[191] Among the few exceptions is Maddy's [1990] rescue effort.

is described in gross, functional terms. These in turn are decomposed into computationally simpler functions. And those are decomposed further. Infinite regress is forestalled by the existence of a basic level of functional decomposition—the *functional architecture* (FA)—which implements directly in the hardware. The basic processes of the functional architecture (whatever they happen to be) are understood, in other words, to be carried out immediately by the physical machine that implements the cognitive process; in the case of artificial models, that can be the virtual machine on which the process is running or perhaps even the machine code itself; in the case of human beings, it's the system of the simplest possible representations that a cognitive process makes use of. More primitive processes are not computational, hence non-cognitive; they are mere physical transitions upon which the FA supervenes. More elaborate cognitive processes, by contrast, employ the structured representations that the FA itself makes available. If we take the computer metaphor seriously, the FA is the canonical language into which all higher level structures can (in principle) be translated. It's that layer of a cognitive system which remains fixed relative its higher-level, computational states. Let me propose that, in modern terms, the reproductive imagination corresponds to the *functional architecture* of the cognitive apparatus—or, at the very least, the functional architecture of the mathematics organ. (Even if this is correct, it helps us but little. Pylyshyn [1984] argues that human brains use functional primitives unlike those used by von Neumann machines. At the moment we have only a very fragmentary understanding of the FA.)[192]

Desiderata. One indication that this way of looking at things is on the right track is that it helps some of our earlier desiderata fall into place. We can borrow our account of the necessity of mathematical content directly from Kant. We have discussed this at length already, so I will be brief. Since the FA is the canonical 'code' in which all higher level representations are constructed, its constraints are inviolable (at least as far humanly possible experience is concerned). We are literally incapable of violating its constraints (though we can misrepresent what those constraints are). Admittedly, this Kantian notion of necessity does not go as deep as some might like. It tells us nothing, for instance, about whether alternative architectures more powerful than our own are possible. Still, this sort of necessity may well be all that finite beings such as ourselves are likely to have access to.[193]

We can likewise borrow our account of the applicability of mathematical representations directly from Kant. The problem, as we saw in Chapter 1, is really twofold. On the one hand, mathematical concepts are indispensable to our *descriptions* of natural phenomena. Indeed, when we are at our most precise, when we are trying to offer a maximally accurate

[192]Again, Biederman [1995] proposes a functional architecture for human vision. Generative grammarians are working toward a theory of universal grammar. Jackendoff [2002] offers us a proposal concerning the basic structures at the semantic interface.

[193]Let me speculate here that if the basic operations of the FA are grounded in fundamental mathematical relations, other intelligent species operate with a relevantly similar fundamental architecture.

description of a natural system, we resort to purely formal notions and mathematical notation. The second (and, if anything, more interesting) feature of mathematical judgements is their uncanny capacity to guide and sometimes even to predict the course of scientific enquiry. Gödel's theory permits us now to start to formulate an explanation of both of these features of mathematical concepts. Mathematics offers us the finest, most powerful language in which to describe natural phenomena since the functional architecture is the finest-grained representational system that our cognitive architecture employs. Little wonder then that it's indispensable to our scientific descriptions of the phenomena. Maths' predictive power is a separate issue. Recall that Gödel suggests that the reproductive imagination is (among other things) a brute, physical, mind-independent phenomenon. Arguably, by developing representations in it, we develop representations of some relatively basic spatio-temporal relations and symmetries. Let me suggest that an isomorphism between local states of the architecture and objective reality underwrites abductive applicability. In effect, we are capable of grasping those aspects of noumenal reality that themselves structure our architecture and that we have learned to represent conceptually. Evidently, this is a speculative solution to our problem. But it has the advantage of being compatible with ontological naturalism (though not with naive realism) and of being substantially different from the platonist resolution of this same issue.

Knowledge. One half of the Benacerraf puzzle was to offer an account of our mathematical knowledge. Let me leave off by making explicit how the ideas discussed here help the epistemologist. Of course, it would be excessively optimistic to expect Gödel's work to resolve the multitude of epistemic problems that surround mathematical knowledge. Nonetheless, I want to suggest that thanks to Gödel we gain a sense of how mathematical beliefs might be justified. To wit, our mathematical knowledge is not any more mysterious than our knowledge of natural language or of slow, medium-sized objects—though it *is* different in kind.

Let's ground the discussion by adopting a working characterization of knowledge, one slightly more nuanced than our earlier attempt. It seems plausible that an agent S knows that a state of affairs p obtains if the following conditions are met:

(1) p;
(2) S explicitly registers (or judges, or represents in an appropriate way) that p; and
(3) S is justified in registering (believing or representing) that p.

Of the three conditions, the first requires the least commentary. Since 'to know' is factive, the condition is uncontentious. The second condition, likewise, has been thoroughly discussed at this point. What is at issue are explicit conceptual structures in the parallel architecture or *the intellect*. The most fraught part of the definition is surely its third plank. Gettier

[1963] cases can be read as showing that our commonsense notion of justification readily lands us in trouble and that a better philosophical account is needed. I want to suggest that (much like our earlier characterization of *meaning*) a theory of *justification* needs to be both descriptively and explanatorily satisfying. That is, it needs to offer a conception of justification that is subtle enough to navigate past Gettier-style scenarios by accurately separating cases of true knowledge from cases of epistemic luck. At the same time, ideally, our account should serve to explain the underlying nature of justification in non-ciruclar, non-epistemic terms that shed light on why some beliefs count as justified while others do not.

The account that best meets these desiderata is, I think, Alvin Goldman's [1979, 1986] reliabilism.[194] The core insight is that the degree of justification enjoyed by a (token) belief varies in proportion to the reliability of the cognitive process (type) that gave rise to it. If S's belief that p's are q's came about as the result of a process that would likely have generated that belief regardless of the state of the p's then S's belief does not count as knowledge. By contrast, if S came to believe that p's are q's via a process that almost certainly would have generated no such belief unless p's really *were* q's then we say that S's belief is justified. (Relying on reliability to account for justification links justification and truth quite closely. But not too closely: S's belief can be justified *and false* if the process by which her belief is formed is not perfectly reliable. That's just as it should be.)

While Goldman's conception of justification sounds sensible enough, it departs from our everyday notion in some important respects that it would be useful to make explicit. In the common parlance, 'justification' functions as a near-synonym for 'evidence'. One's justification for an opinion or position that one cleaves to is precisely the sort of thing that one can be expected to report on demand. At the very least, it's the sort of thing that one ought to be able, in principle, to bring to consciousness. The technical notion of justification elaborated by Goldman departs from this usage. A cognitive process can be reliable even if it's irretrievably unconscious and incapable of being clearly conceptualized. In effect, one can hold a justified belief but be wholly incapable of articulating reasons for holding it. In fact, an epistemic agent can hold a justified belief even if she is unaware that she holds it. Again, *all* that is required for a belief to be justified—and hence potentially to count as knowledge—is for it to have been formed by internal, cognitive processes that typically gives rise to error-free representations of some relevant facts.[195]

There is one additional aspect of the theory that it will be helpful for us to make explicit. It's possible to distinguish two broad kinds of cognitive processes relevant to this discussion. The first sort are reliable (or unreliable) regardless of the particular beliefs (or desires) the agent happens to have. Simple perceptual beliefs are of this sort. Visual illusions affect us

[194]I confine myself here to Goldman's reliabilism about justification.

[195]Notice that Goldman's justification is not an all or nothing affair. Compare DeRose [2002].

regardless of our artistic preferences or political beliefs. Indeed, they affect us even if we know *that* we are being subjected to an illusion and how the illusion operates. Likewise, veridical perception is epistemically reliable in the relevant sense even if we are sometimes prone not to trust our eyes or actively misinterpret what we see. There is however another sort of reliable cognitive process. It too is capable of giving rise to knowledge. This second type of process can result in knowledge only on condition that the beliefs that serve as (some of) its inputs are themselves true. Valid logical arguments are reliable in just this sense. And so are mathematical proofs predicated on the correctness of some antecedently assumed point of departure. The reliability of the conclusions arrived at depends on the structure of the arguments *and* the truth of the premises with which we start.

Goldman [1979] proposes a definition of justified belief that distinguishes these two cases:

(3.a) If S's belief in p at t results ('immediately') from a belief-independent process that is unconditionally reliable, then S's belief in p at t is justified.

(3.b) If S's belief in p at t results ('immediately') from a belief-dependent process that is (at least) conditionally reliable, and if the beliefs (if any) on which the process operates in producing S's belief in p at t are themselves justified, then S's belief in p at t is justified.

(3.c) No other beliefs are justified.

To sum up: Belief-independent doxogenetic processes are justified to the extent that they themselves are reliable. Belief-dependent (conditional) representations are reliable so long as at least one of the causal paths leading to their formation is itself reliable, but not otherwise. Nothing else counts as justified.[196]

A belief in some factual state of affairs p is justified in case it results ('immediately') from a belief-independent process that is unconditionally reliable. Many perceptual processes are unconditionally reliable in this sense; they are not dependent on our beliefs, desires or other propositional attitudes. (Visual illusions are the obvious illustration.) Moreover, visual perception is liable to offer accurate information. Now, by construing the content of fundamental mathematical judgements as conceptually representing the states of the FA (in some to be specified sense) we can take advantage of this same account.[197] After all, under one aspect, the FA is a brute physical process; it's no less objective than physical processes

[196]It has been noted (*inter alia* by Nozick [1981]) that mathematical knowledge constitutes one of the hard cases for the reliabilist. That's good news however insofar as it means that we don't stack the deck by adopting reliabilism.

[197]This does *not* mean however that mathematical concepts are *about* the FA! They are, evidently, about the extensions of mathematical concepts: the mathematical entities that we experience ourselves as manipulating. It's important to keep track of which of Sellars' stereoscopic images we are working with.

taking place at a remove from the body. Perceptual error results from mis-attributions of properties. Similarly, mathematical error derives from a mis-representation of the states of the architecture. We do not thereby solve the problem of error but we do show that the problem of mathematical error is solvable *modulo* a description of the fundamental representations used at FA and a solution to the problem of perceptual error more generally. Adopting Gödel's account we therefore gain an account of the necessity of mathematical propositions, an account of their applicability, and an indication of how to go about constructing a philosophical theory of mathematical knowledge. These are, I think, nontrivial gains. So even though we are still some distance from achieving our scientific goals, at least some of our philosophical worries are starting to be addressed.

Conclusion

I have attempted to offer a hypothesis concerning the source of constraints operative on our semantic representations of mathematical content. I began by arguing that current neuroscientific and psychological theories, while helpful, can only be a part of the story. What we ultimately need is an account of the basic representations that underwrite mathematical cognition as well as of how they are recombined in mathematical thought. Further, we need an account of why these representations are special: why they track facts that are deep enough so as not to be contingent and fine enough so as to help us make finer cuts than natural language can. In effect, we need a philosophically responsible, algorithmic-level theory of structure for the mathematics faculty (perhaps of the sort now available for language and for vision). In the second half of the chapter, I suggested that the beginnings of the sort of theory we need can be found in the work of Kurt Gödel. I finished by suggesting that Gödel's work can be interpreted in modern terms if we read Kant's reproductive imagination as (in essence) the functional architecture of the mathematics faculty. I suggested also that a reliabilist account of the justification of mathematical judgements is possible.

5 Final Remarks

I study Mathematics as a product
of the human mind, not as absolute.
— Emil Post

The nature of mathematical reality has been a thorn in the side of ontological naturalism for some time. Some philosophers have resigned themselves to counting acausal, abstrct objects among legitimate naturalist posits.[198] Alan Weir [2005] summarizes the current situation with these words:

> On the face of it, mathematics is an enormous Trojan Horse sitting firmly in the centre of the citadel of naturalism. Modern natural science is mathematical through and through: it is impossible to do physics, chemistry, molecular biology and so forth without a very thorough and quite extensive knowledge of modern mathematics (indeed this is true to an increasing extent of social sciences such as psychology and economics). Yet, prima facie, mathematics provides a counter-example... to ontological naturalism.

Among my principal aims in this dissertation has been to disarm the threat to ontological naturalism posed by acausal abstracta. I began by offering arguments for the reality of mathematical entities based on math's usefulness to deductive reasoning in the sciences, as well as its unique role in the construction of creative analogies that push forward scientific discovery. These two aspects of math's indispensability to natural science should not, I think, be ignored. They both require explanation and, in my view, only an ontologically realist conception of mathematics has much hope of offering one. (Some may disagree.) In Chapter 2, I argued that realism about *acausal* entities is a dead end. It does not help explain our access to mathematical facts and even our best stabs at how knowledge of an acausal domain might be achieved run counter to available empirical evidence. One of the reasons, I think, that many researchers have felt the need to commit to abstract objects is a problematic view of natural language semantics. It has seemed sensible to assume that the truth of such statements as that France is a hexagonal republic requires that there exist a freestanding entity (France) with certain further properties. Things are not quite so straightforward. In Chapter 3, I reviewed arguments showing that traditional extensional semantics is wrongheaded; seemingly fundamental word-world relations draw inexorably on human perceptual and conceptual capacities. The character of the Benacerraf puzzle alters considerably once we commit to an internalist, mentalist conception of meaning. In particular, rather than trying to discover how we glean truths about otherworldly mathematical entities, we are called upon the explain how the concepts that underwrite our mathematical reasoning are constrained. In Chapter 4, I explored the possibility that the epistemic

[198] Avowedly supernaturalist philosophers, including Jerrold Katz [1990] and Mark Steiner [2005], have been more than happy to highlight the cognitive dissonance engendered by this balancing act.

constraints on mathematical concepts derive from the structure of the human cognitive apparatus. On this account, due originally to Gödel's reading of Kant, mathematical content is ultimately grounded in the spatio-temporal structure of our functional architecture. The subtle mathematical ideas manipulated by the professional involve the construction of various complexes, elaborations, and meta-representations of that basic structure.

I'd like to end by emphasizing a corollary of the view advanced here. Namely this: If constraints on mathematical judgements derive directly from the structure of our transcendental functional architecture then the structure of our transcendental functional architecture can be recognized in what we have been calling the "mathematical landscape." The foundations of mathematics constitute an important source of data about the nature of the mind. The (non-trivial) challenge facing us is to learn to systematize and interpret this data. It's to be hoped that by studying the foundations of mathematics we may perhaps also learn a little about the very basic cognitive operations that make us what we are.

References

Albert, M. S., Diamond, A. D., Fitch, R. H., Neville, H. J., Rapp, P. R., and Tallal, P. A. (1999). Cognitive development. In Zigmond, M. J., Bloom, F. E., Landis, S. C., Roberts, J. L., and Squire, L. R., editors, *Fundamental Neuroscience*. Academic Press.

Anderson, M., O'Connor, N., and Hermelin, B. (1999). A specific calculating ability. *Intelligence*, 26(4):383–403.

Aristotle (1941). *Basic Works*. Random House.

Balaguer, M. (1998). *Platonism and anti-platonism in mathematics*. Oxford University Press.

Barnett, R. M., Mühry, H., and Quinn, H. R. (2000). *The Charm of Strange Quarks*. Springer-Verlag.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22:577–609.

Bell, J. L. (2008/1988). *Toposes and Local Set Theories*. Dover.

Bell, J. L., DeVidi, D., and Solomon, G. (2001). *Logical Options: An Introduction to Classical and Alternative Logics*. Broadview Press.

Benacerraf, P. (1965). What numbers could not be. In Benacerraf, P. and Putnam, H., editors, *Philosophy of Mathematics: Selected Readings*. Cambridge University Press.

Benacerraf, P. (1973). Mathematical truth. In Benacerraf, P. and Putnam, H., editors, *Philosophy of Mathematics: Selected Readings*. Cambridge University Press.

Bendegem, J. P. V. and Kerkhove, B. V. (2004). The unreasonable richness of mathematics. *Journal of Cognition and Culture*, 4.3:525–549.

Berthier, M., Starkstein, S., Lylyk, P., and Leiguardia, R. (1990). Differential recovery of languages in a bilingual patient: a case study using selective amytal test. *Brain and Language*, 38(449-453).

Biederman, I. (1995). Visual object recognition. In Kosslyn, S. M. and Osherson, D. N., editors, *Visual Cognition*, An Invitation to Cognitive Science, pages 121–166. MIT Press.

Bijeljac-Babic, R., Bertoncini, J., and Mehler, J. (1993). How do four-day-old infants categorize multisyllabic utterances? *Developmental Psychology*, 29(4):711–721.

Bishop, A. J. (2002). Mathematical acculturation, cultural conflicts, and transition. In de Abreu, G., J.Bishop, A., and Presmeg, N. C., editors, *Transitions between contexts of mathematical practices*, pages 193–212. Kluwer Academic Press.

Brannon, E. M. (2005). The independence of language and mathematical reasoning. *Proceedings of the National Academy of Science (USA)*, 102(9).

Brook, A. (1994). *Kant and the Mind*. Cambridge University Press.

Brook, A. (2004). Kant, cognitive science, and contemporary neo-Kantianism. *Journal of Consciousness Studies*, 11(10–11):1–25.

Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence*, 47:139–159.

Brouwer, L. E. J. (1975). *L. E. J. Brouwer: Collected Works*, volume 1. North-Holland Publishing Co.

Brown, J. R. (1999). *Philosophy of Mathematics: An Introduction to the World of Proofs and Pictures*. Routledge.

Buck-Morss, S. (1991). *The Dialectics of Seeing: Walter Benjamin and the Arcades Project*. MIT Press.

Burgess, J. P. and Rosen, G. (1997). *A Subject with No Object: Strategies for Nominalistic Interpretations of Mathematics*. Oxford University Press.

Butterworth, B. (1999). *What Counts: How Every Brain is Hardwired for Math*. Free Press.

Butterworth, B. (2005). Developmental dyscalculia. In Campbell, J. I., editor, *Handbook of Mathematical Cognition*, pages 455–467. Psychology Press.

Campbell, J. I., editor (2005). *Handbook of Mathematical Cognition*. Taylor & Francis.

Cann, R. (1993). *Formal Semantics: An Introduction*. Cambridge University Press.

Cappelletti, M., Butterworth, B., and Kopelman, M. (2001). Spared numerical abilities in a case of semantic dementia. *Neuropsychologia*, 39(11):1224–1239.

Carey, S. (2001). Cognitive foundations of arithmetic: Evolution and ontogenesis [review of *The Number Sense*]. *Mind and Language*, 16(1):37–55.

Carey, S. (2004). Bootstrapping & the origin of concepts. American Academy of Arts and Sciences.

Carson, E. (1996). *Mathematics, metaphysics and intuition in* Kant. PhD thesis, Harvard University.

Carson, E. (1999). Kant on the method of mathematics. *Journal of the History of Philosophy*, 37(4):629–652.

Carson, E. (2004). Metaphysics, mathematics and the distinction between the sensible and the intelligible in Kant's Inaugural Dissertation. *Journal of the History of Philosophy*, 42(2):165–194.

Changeux, J.-P. and Connes, A. (1995). *Conversations on Mind, Matter, and Mathematics*. Princeton University Press.

Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT Press.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin and Use*. Convergence.

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge University Press.

Churchland, P. S. (1998). Brainshy: Nonneural theories of conscious experience. In Hameroff, S., Kaszniak, A., and Laukes, J., editors, *Toward a Science of Consciousness II: The second Tucson discussions and debates*, pages 109–126. MIT Press.

Churchland, P. S., Ramachandran, V., and Sejnowski, T. J. (1994). A critique of pure vision. In Koch, C. and Davis, J. L., editors, *Large-Scale Neuronal Theories of the Brain*. MIT Press.

Churchland, P. S. and Sejnowski, T. J. (1992). *The Computational Brain*. MIT Press.

Cipolotti, L., Butterworth, B., and Denes, G. (1991). A specific deficit for numbers in a case of dense acalculia. *Brain*, 114:2619–2637.

Colyvan, M. (2001). *The Indispensability of Mathematics*. Oxford University Press.

Cook, V. J. and Newson, M. (2007). *Chomsky's Universal Grammar*. Blackwell Publishers, 3rd. edition.

Corry, L. (1998). The influence of David Hilbert and Herman Minkowski on Einsten's views over the interrelation between physics and mathematics. *Endeavour*, 22(3):95–7.

Culicover, P. W. and Jackendoff, R. (2005). *Simpler Syntax*. Oxford University Press.

Cytowic, R. E. (2003). *The man who tasted shapes*. MIT Press.

Das, A. and Ferbel, T. (2003). *Introduction to Nuclear and Particle Physics*. World Scientific, 2nd edition.

Davidson, D. (1967). Truth and meaning. In Martinich, A. P., editor, *The Philosophy of Language*, chapter 6, pages 79–90. Oxford University Press.

Davis, P. J. and Hersh, R. (1998). *The Mathematical Experience*. Mariner Books.

Dawson, J. W. (1997). *Logical Dilemmas: The Life and Work of Kurt Gödel*. A K Peters.

Dawson, M. (1999). *Understanding Cognitive Science*. Oxford University Press.

Dehaene, S. (1997). *The Number Sense: How the Mind Creates Mathematics*. Oxford University Press.

Dehaene, S. (2003). The neural basis of the Weber-Fechner law: a logarithmic mental number line. *Trends in Cognitive Sciences*, 7(4).

Dehaene, S., Piazza, M., Pinel, P., and Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, 20:487–506.

DeLong, H. (2004). *A Profile of Mathematical Logic*. Dover.

Dennett, D. C. (1987). *The Intentional Stance*. MIT Press.

Dennett, D. C. (1991). *Consciousness Explained*. Back Bay Books.

Dennett, D. C. (1998). True believers. In *The Intentional Stance*. MIT Press.

DeRose, K. (2002). Assertion, knowledge, and context. *Philosophical Review*.

Detlefsen, M. (2005). Formalism. In Shapiro, S., editor, *Oxford Hanbook of the Philosophy of Mathematics and Logic*. Oxford University Press.

Donlan, C., Cowan, R., Newton, E. J., and Lloyd, D. (2007). The role of language in mathematical development: Evidence from children with specific language impairments. *Cognition*, 103:23–33.

Duhem, P. (1906). *The Aim and Structure of Physical Theory*. Princeton University Press, 1954 edition.

Duhem, P. (1915). *La Science allemande*. A. Hermann et Fils.

Dummett, M. (1994). What is mathematics about? In George, A., editor, *Mathematics and Mind*. Oxford University Press.

Einstein, A. (1961). *Relativity: The Special and the General Theory*. Three Rivers Press.

Enderton, H. B. (1977). *Elements of Set Theory*. Academic Press.

Euclid (1956). *The Thirteen Books of Euclid's Elements*. Dover, 1956 edition.

Eves, H. (1958/1997). *Foundations and Fundamental Concepts of Mathematics*. Dover, 3rd. edition.

Falkenstein, L. (1990). Was Kant a nativist? *Journal of the History of Ideas*, 51(4):573–597.

Falkenstein, L. (1995). *Kant's Intuitionism: A Commentary on the Transcendental Aesthetic*. University of Toronto Press.

Fauconnier, G. (1994). *Mental Spaces*. Cambridge University Press.

Field, H. (1989). *Realism, Mathematics and Modality*. Blackwell.

Fodor, J. (1975). *The Language of Thought*. Harvard University Press.

Fodor, J. (1987). *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. MIT Press.

Franks, N. P. and Lieb, W. R. (1998). The molecular basis of general anesthesia: Current ideas. In Hameroff, S., Kaszniak, A., and Laukes, J., editors, *Toward a Science of Consciousness II: The second Tucson discussions and debates*, pages 444–455. MIT Press.

Frege, G. (1953). *The Foundations of Arithmetic*. Blackwell.

French, S. (2000). The reasonable effectiveness of mathematics: partial structures and the application of group theory to physics. *Synthese*, 125:103–120.

Frith, U. and Happé, F. (1999). Theory of mind and self-consciousness: What is it like to be autistic? *Mind and Language*, 14(1):1–22.

Gallistel, C. R. and Gelman, R. (2000). Non-verbal numerical cognition: from reals to integers. *Trends in Cognitive Sciences*, 4(2).

Gallistel, C. R., Gelman, R., and Cordes, S. (2005). The cultural and evolutionary history of the real numbers. In *Culture and evolution*, Fryssen Foundation. MIT Press.

Gelman, R. and Butterworth, B. (2005). Number and language: how are they related? *Trends in Cognitive Sciences*, 9(1).

Gelman, R. and Gallistel, C. R. (1978). *The Child's Understanding of Number*. Harvard University Press.

George, A. and Velleman, D. J. (2002). *Philosophies of Mathematics*. Blackwell.

Gettier, E. L. (1963). Is justified true belief knowledge? *Analysis*, 23:121–123.

Giaquinto, M. (2001). What cognitive systems underlie arithmetical abilities? *Mind and Language*, 16(1):56–68.

Gillies, D. (1998). The Duhem thesis and the Quine thesis. In Curd, M. and Cover, J. A., editors, *Philosophy of Science*. Norton.

Gleitman, L. and Fisher, C. (2005). Universal aspects of word learning. In McGilvray, J., editor, *Chomsky*. Cambridge University Press.

Gödel, K. (1947). What is Cantor's continuum problem? In Feferman, S., Dawson, J. W., Kleene, S., Moore, G., Solovay, R., and van Heijenoort, J., editors, *Kurt Gödel: Collected Works*, volume II, pages 176–187. Oxford University Press.

Gödel, K. (1949). Some observations about the relationship between theory of relativity and Kantian philosophy. In Feferman, S., Dawson, J. W., Goldfarb, W., Parsons, C., and Solovay, R. M., editors, *Kurt Gödel: Collected Works*, volume III, pages 230–246. Oxford University Press.

Gödel, K. (1961). The modern development of the foundations of mathematics in the light of philosophy. In Feferman, S., Dawson, J. W., Goldfarb, W., Parsons, C., and Solovay, R. M., editors, *Kurt Gödel: Collected Works*, volume III, pages 375–387. Oxford University Press.

Goldman, A. I. (1967). A causal theory of knowing. *The Journal of Philosophy*, 64(12):357–372.

Goldman, A. I. (1979). What is justified belief? In Pappas, G. S., editor, *Justification and Knowledge*, pages 1–23. Reidel Publishing Company.

Goldman, A. I. (1986). *Epistemology and Cognition*. Harvard University Press.

Goodale, M. A. (1995). The cortical organization of visual and visumotor control. In Osherson, D. N., editor, *An Invitation to Cognitive Science: Visual Cognition*. MIT Press, 2nd edition.

Gordon, P. (2004). Numerical cognition without words: evidence from Amazonia. *Science*.

Gurney, K. (1997). *An Introduction to Neural Networks*. Routledge.

Haack, S. (1978). *Philosophy of Logics*. Cambridge University Press.

Hameroff, S. R. (1998). "More Neural Than Thou". In Hameroff, S., Kaszniak, A., and Laukes, J., editors, *Toward a Science of Consciousness II: The second Tucson discussions and debates*, pages 197–213. MIT Press.

Hanson, N. R. (1963). *The Concept of the Positron.* Cambridge University Press.

Hardy, G. H. (1940). *A Mathematician's Apology.* University of Alberta Mathematical Sciences Society.

Hauser, M. D., Chomsky, N., and Fitch, W. T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science*, 298.

Hauser, M. D. and Spelke, E. S. (2004). Evolutionary and developmental foundations of human knowledge. In Gazzaniga, M., editor, *The Cognitive Neurosciences, III.* MIT Press.

Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar.* Blackwell Publishers.

Hellman, G. (1989). *Mathematics without Numbers: Towards a Modal-Structural Interpretation.* Clarendon Press.

Hellman, G. (2001). Three varieties of mathematical structuralism. *Philosophia Mathematica*, 9(3):184–211.

Hellman, G. (2003). Does category theory provide a framework for mathematical structuralism? *Philosophia Mathematica*, 11:129–157.

Hempel, C. G. (1998/1962). Two basic types of scientific explanation. In Curd, M. and Cover, J. A., editors, *Philosophy of Science.* Norton.

Hermelin, B. (2001). *Bright Splinters of the Mind.* Jessica Kingsley Publishers.

Hermelin, B. and O'Connor, N. (1983). The idiot-savant: flawed genius or clever Hans. *Psychological Medicine*, 13:479–481.

Hermelin, B. and O'Connor, N. (1990). Factors and primes: a specific numerical ability. *Psychological Medicine*, 20:163–169.

Heyting, A. (1956). *Intuitionism: An Introduction.* North-Holland Publishing Co.

Hinzen, W. (2006). *Mind Design and Minimalist Syntax.* Oxford University Press.

Hoffman, S. (1999). *Mathematics as make-believe: a constructive empiricist account.* PhD thesis, University of Alberta.

Horwich, P. (1997). Implicit definition, analytic truth, and apriori knowledge. *Nous*, 31(4):423–440.

Houlgate, S. (2006). *The opening of Hegel's logic: from being to infinity.* Purdue University Press, 2006.

Hume, D. (1993/1748). *An Enquiry Concerning Human Understanding*. Hackett.

Ingram, J. C. L. (2007). *Neurolinguistics: An introduction to spoken language processing and its disorders*. Cambridge Textbooks in Linguistics. Cambridge University Press.

Jackendoff, R. (1983). *Semantics and Cognition*. MIT Press.

Jackendoff, R. (1987). *Consciousness and the Computational Mind*. Bradford MIT Press.

Jackendoff, R. (1992a). Languages of the mind. In *Languages of the Mind*. MIT Press.

Jackendoff, R. (1992b). What is a concept, that a person may grasp it? In *Languages of the Mind*. MIT Press.

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford University Press.

Jackendoff, R. (2003). Précis of the Foundations of Language. *Behavioral and Brain Sciences*, 26(6).

Jackendoff, R. (2007). A parallel architecture perspective on language processing. *Brain Research*, pages 2–22.

Jensen, A. R. (1979). Reaction time, movement time, and intelligence. *Intelligence*, 3:121–126.

Jerzykiewicz, L. and Scott, S. (2003). Psychologism and conceptual semantics: Review of *Foundations of Language*. *Behavioral and Brain Sciences*, 26(6):682–683.

Kadosh, R. C., Kadosh, K. C., Schuhmann, T., Kaas, A., Goebel, R., Henik, A., and Sack, A. T. (2007). Virtual dyscalculia induced by parietal-lobe tms impairs automatic magnitude processing. *Current Biology*, 17:1–5.

Kager, R. (1999). *Optimality Theory*. Cambridge University Press.

Kant, I. (1996/1781). *Critique of Pure Reason*. Hackett.

Kant, I. (1997/1783). *Prolegomena to Any Future Metaphysics*. Cambridge University Press.

Katz, J. J. (1990). *The Metaphysics of Meaning*. MIT Press.

Katz, J. J. (1998). *Realistic Rationalism*. MIT Press.

Kawai, N. and Matsuzawa, T. (2000). Numerical memory span in a chimpanzee. *Nature*, 403:39–40.

Kitcher, P. (1983). *The Nature of Mathematical Knowledge*. Oxford University Press.

Kitcher, P. (1990). *Kant's Transcendental Psychology*. Oxford University Press.

Kline, M. (1972). *Mathematical Thought from Ancient to Modern Times*. Oxford University Press.

Kline, M. (1980). *Mathematics: The loss of certainty*. Oxford University Press.

Koch, C. (2004). *The Quest for Consciousness*. Roberts and Company.

Kuhn, T. S. (1970). *The Structure of Scientific Revolutions*. University of Chicago Press.

Kusch, M. (1995). *Psychologism: a case study in the sociology of philosophical knowledge*. Routledge.

Landau, B., Gleitman, H., and Spelke, E. (1981). Spatial knowledge and geometric representation in a child blind from birth. *Science*, 213(4513):1275–1278.

Landau, B., Spelke, E., and Gleitman, H. (1984). Spatial knowledge in a young blind child. *Cognition*, 16:225–260.

Landerl, K., Bevana, A., and Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8-9-year-old students. *Cognition*, 93.

Larson, R. and Segal, G. (1995). *Knowledge of Meaning: An Introduction to Semantic Theory*. MIT Press.

Laudan, L. (1990). Demystifying underdetermination. In Savage, W., editor, *Minnesota Studies in the Philosophy of Science*. University of Minnesota Press.

Laurence, S. and Margolis, E. (2005). Number and natural language. In Carruthers, P., Laurence, S., and Stich, S., editors, *The Innate Mind: Structure and Content*, pages 216–238.

Lecours, A. R. and Joanette, Y. (1980). Linguistic and other psychological aspects of paroxysmal aphasia. *Brain and Language*, 10:1–23.

Lee, S. A. and Spelke, E. S. (2008). Children's use of geometry for reorientation. *Developmental Science,*, 11(5):743–749.

Liebeck, M. W. (2008). The insolubiliy of the quintic. In Gowers, T., editor, *The Princeton Companion to Mathematics*, pages 708–710. Princeton University Press.

Mac Lane, S. (1986). *Mathematics Form and Function*. Springer-Verlag.

Mac Lane, S. (1997). Despite physics, proof is essential in mathematics. *Synthese*, 111:147–154.

Maddy, P. (1990). *Realism in Mathematics*. Oxford University Press.

Maddy, P. (1997). *Naturalism in Mathematics*. Oxford University Press.

Maddy, P. (2007). *Second Philosophy: A Naturalistic Method*. Oxford University Press.

Margolis, E. and Laurence, S. (1999). *Concepts: Core Readings*. MIT Press.

Marr, D. (1982). *Vison*. Freeman.

McGilvray, J. (1998). Meanings are syntactically individuated and found in the head. *Mind and Language*, 13(2):225–280.

McGilvray, J. (2005). *The Cambridge Companion to Chomsky*. Cambridge University Press.

McGinn, C. (2000). *Logical Properties*. Oxford University Press.

Meck, W. H. and Church, R. M. (1983). A mode control model of counting and timing processes. *Journal of Experimental Psychology: Animal Behavior Processes*, 9(320-334).

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge University Press.

Miller, K. F., Kelly, M., and Zhou, X. (2005). Learning mathematics in China and the United States: Cross-cultural insights into the nature and course of preschool mathematical development. In Campbell, J. I., editor, *Handbook of Mathematical Cognition*, pages 163–178. Psychology Press.

Montague, R. (1974a). English as a formal language. In Thomason, R. H., editor, *Formal Philosophy: Selected Papers of Richard Montague*, pages 188–221. Yale University Press.

Montague, R. (1974b). *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press.

Montague, R. (1974c). Universal grammar. In Thomason, R. H., editor, *Formal Philosophy: Selected Papers of Richard Montague*, pages 222–246. Yale University Press.

Nagel, E. and Newman, J. R. (2001). *Gödel's Proof*. New York University Press.

Nozick, R. (1981). *Philosophical Explanations*. Belknap Press.

O'Reilly, R. C. and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press.

Ouhalla, J. (1994). *Introducing Transformational Grammar: From Principles and Parameters to Minimalism*. Arnold Press, 2nd edition.

Papineau, D. (2001). The rise of physicalism. In Gillett, C. and Loewer, B., editors, *Physicalism and Its Discontents*, pages 3–36. Cambridge University Press.

Parshall, K. H. (2008). The development of abstract algebra. In Gowers, T., editor, *The Princeton Companion to Mathematics*, pages 95–106. Princeton University Press.

Parsons, C. (1975). Quine on the philosophy of mathematics. In *Mathematics in Philosophy*. Cornell University Press.

Parsons, C. (1983). *Mathematics in Philosophy*. Cornell University Press.

Pascalis, O., de Schonen, S., Morton, J., Deruelle, C., and Fabre-Grenet, M. (1995). Mother's face recognition by neonates: a replication and an extension. *Infant behavior and development*, 18(1):79–85.

Penrose, R. (1989). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford University Press.

Petitto, L.-A. (2005). How the brain begets language. In McGilvray, J., editor, *The Cambridge Companion to Chomsky*. Cambridge University Press.

Pietroski, P. M. (2003a). The character of natural language semantics. In Barber, A., editor, *Epistemology of Language*. Oxford University Press.

Pietroski, P. M. (2003b). Meaning before truth.

Pietroski, P. M. and Crain, S. (2002). Why language acquisition is a snap. *Linguistic Review*, 19:163–83.

Pietroski, P. M. and Crain, S. (2005). Innate ideas. In McGilvray, J., editor, *The Cambridge Companion to Chomsky*. Cambridge University Press.

Pinker, S. (1989). *Learnability and Cognition: The Acquisition of Argument Structure*. MIT Press.

Popper, K. R. (1963). *Conjectures and Refutations: The growth of scientific knowledge*. Routledge, 1996 edition.

Posy, C. (2005). Intuitionism and philosophy. In Shapiro, S., editor, *Oxford Hanbook of the Philosophy of Mathematics and Logic*, pages 318–355. Oxford University Press.

Pulvermüller, F. (2002). *The Neuroscience of Language: On Brain Circuits of Words and Serial Order*. Cambridge University Press.

Putnam, H. (1967). Mathematics without foundations. In Benacerraf, P. and Putnam, H., editors, *Philosophy of Mathematics: Selected Readings*, pages 295–311. Cambridge University Press.

Putnam, H. (2004). *Ethics without ontology.* Harvard University Press.

Putnam, H. (2006). Indispensability arguments in the philosophy of mathematics. The 40^{th} Chapel Hill Colloquium in Philosophy.

Pyenson, L. (1977). Hermann Minkowski and Einstein's Special Theory of Relativity: With an appendix of Minkowski's *Funktiontheorie* manuscript. *Archive for History of Exact Sciences*, 17(1):71–95.

Pylyshyn, Z. W. (1984). *Computation and Cognition.* MIT Press.

Pylyshyn, Z. W. (2001). Visual indexes, preconceptual objects, and situated vision. *Cognition*, 80:127–158.

Quine, W. V. O. (1953). Two dogmas of empiricism. In *From a Logical Point of View*, pages 20–46. Harvard University Press.

Ravizzaa, S. M., Andersonb, J. R., and Carter, C. S. (2008). Errors of mathematical processing: The relationship of accuracy to neural regions associated with retrieval or representation of the problem state. *Brain Research*, (1238):118–126.

Regier, T. (1996). *The Human Semantic Potential: Spatial Language and Constrained Connectionism.* MIT Press.

Resnik, M. D. (1997). *Mathematics as a Science of Patterns.* Oxford University Press.

Riordan, M. (1987). *The Hunting of the Quark.* Simon & Schuster.

Rosen, G. (2001). Nominalism, naturalism, epistemic relativism. *Philosophical Perspectives*, 15.

Ryle, G. (1949). *The Concept of Mind.* University of Chicago Press, 1984 edition.

Sakezles, P. (2009). The Aristotelian origins of Stoic determinism. *Proceedings of the Boston Area Colloquium in Ancient Philosophy*, (24).

Scott, S. (2003). *Non-referring concepts.* PhD thesis, Carleton University.

Sellars, W. (1962). Philosophy and the scientific image of man. In *Frontiers of Science and Philosophy*, pages 35–78. University of Pittsburgh Press.

Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.

Shapiro, S. (2000). *Thinking about Mathematics: The Philosophy of Mathematics*. Oxford University Press.

Shelton, P. A., Bowers, D., Duara, R., and Heilman, K. M. (1994). Apperceptive visual agnosia: A case study. *Brain and Cognition*, 25:1–23.

Skyrms, B. (1967). The explication of "X knows that p". *The Journal of Philosophy*, 64(12):373–389.

Smith, N. (1999). *Chomsky: Ideas and Ideals*. Cambridge University Press.

Solomon, M. K. (2009). On Kurt Gödel's philosophy of mathematics. Technical report, Department of Computer Science and Engineering, Florida Atlantic University.

Spelke, E. S. (2002). What makes us smart? Core knowledge and natural language. In Gentner, D. and Goldin-Meadow, S., editors, *Language in mind: Advances in the study of language and thought*. MIT Press.

Spelke, E. S. and Tsivkin, S. (2001). Initial knowledge and conceptual change: space and number.

Spurrett, D. J. (1999). *The Completeness of Physics*. PhD thesis, University of Natal.

Stainton, R. J. (2006a). Meaning and reference - some Chomskian themes. In Lepore, E. and Smith, B., editors, *Handbook of Philosophy of Language*. Oxford University Press.

Stainton, R. J. (2006b). *Words and Thoughts: Subsentences, Ellipsis, and the Philosophy of Language*. Oxford University Press.

Stanley, J. (2001). Hermeneutic fictionalism. *Midwest Studies in Philosophy*, XXV.

Steiner, M. (1998). *The Applicability of Mathematics as a Philosophical Problem*. Harvard University Press.

Steiner, M. (2005). Mathematics—application and applicability. In Shapiro, S., editor, *Oxford Handbook of Philosophy of Mathematics and Logic*, pages 625–650. Oxford University Press.

Tager-Flusberg, H., Joseph, R., and Folstein, S. (2001). Current directions in research on autism. *Mental Retardation and Developmental Disabilities Research Reviews*, 7:21–29.

Talmy, L. (2001). *Toward a Cognitive Semantics: Volume I: Concept Structuring Systems*. MIT Press.

Tarski, A. (2001). The semantic conception of truth and the foundations of semantics. In Lynch, M. P., editor, *The Nature of Truth*, pages 331–363. MIT.

Thomason, R. H. (1974). Introduction. In *Formal Philosophy: Selected Papers of Richard Montague*. Yale University Press.

Thomason, R. H. (1995). Some limitations to the psychological orientation in semantic theory. Unpublished Draft.

Treisman, A. and Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136.

Trick, L. M. and Pylyshyn, Z. W. (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(2):331–351.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59:433–460.

van Fraassen, B. (1998). Arguments concerning scientific realism. In Cover, J. A. and Curd, M., editors, *Philosophy of Science*. W.W. Norton & Co.

Varley, R. A., Klessinger, N. J. C., Romanowski, C. A. J., and Siegal, M. (2005). Agrammatic but numerate. *Proceedings of the National Academy of Science (USA)*, (pnas.0407470102).

Wang, H. (1987). *Reflections on Kurt Gödel*. MIT Press.

Weir, A. (2005). Naturalism reconsidered. In Shapiro, S., editor, *Oxford Hanbook of the Philosophy of Mathematics and Logic*, pages 460–482. Oxford University Press.

Weyl, H. (2009/1927). Time relations in the cosmos, proper time, lived time, and metaphysical time. In *Mind and Nature: selected writings on philosophy, mathematics, and physics*. Princeton University Press.

Weyl, H. (2009/1934). Mind and nature. In *Mind and Nature: selected writings on philosophy, mathematics, and physics*. Princeton University Press.

Wigner, E. (1960). The unreasonable effectiveness of mathematics in the natural sciences. *Communications in Pure and Applied Mathematics*, 13(1).

Williamson, T. (2004). Must do better. In Greenough, P. and Lynch, M., editors, *St. Andrews Conference on Realism and Truth*. Oxford University Press.

Wynn, K. (1992). Evidence against empiricist accounts of the origins of numerical knowledge. *Mind and Language*, 7(4):315–332.

Wynn, K. (1998). Psychological foundations of number: numerical competence in human infants. *Trends in Cognitive Sciences*, 2(8).

Xu, F. and Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74:B1–11.

Yablo, S. (2001). Go figure: A path through fictionalism. *Midwest Studies in Philosophy*, XXV.

Yamada, J. (1990). *Laura: a case for the modularity of language.* MIT Press.

Yourgrau, P. (2005). *A world without time: the forgotten legacy of Gödel and Einstein.* Basic Books.

Curriculum Vitae

Name:	Luke V. Jerzykiewicz
Post-secondary Education and Degrees:	University of Western Ontario London, Ontario, Canada Ph.D. Philosophy 2005-2009
	Carleton University Ottawa, Ontario, Canada Doctoral Studies in Cognitive Science 1999-2005
	University of Warick Coventry, United Kingdom M.A. Philosophy and Literature 1997-1998
	McGill University Montréal, Québec, Canada B.A. (hon.) Philosophy 1993-1995
	University of King's College Halifax, Nova Scotia, Canada Foundation Year Programme 1992
Honours and Awards:	University of Western Ontario Student Council Teaching Honour Roll 2005-2006, 2003-2004
	Ontario Graduate Scholarship 2003-2005
	Social Science and Humanities Research Council (SSHRC) Doctoral Fellowship 2001-2003

Teaching Experience:

Visiting Assistant Professor
Wellesley College
Wellesley, Massachusetts, USA
2008-2009

Visiting Assistant Professor
Clark University
Worcester, Massachusetts, USA
2007-2008

Visiting Lecturer
Bridgewater State College
Bridgewater, Massachusetts, USA
2006-2007

Lecturer
University of Western Ontario
London, Ontario, Canada
2003-2006

Sessional Lecturer
Carleton University
Ottawa, Ontario, Canada
2003

Publications:

Jerzykiewicz, L. and Scott, S. (2003). Psychologism and Conceptual Semantics, Behavioral and Brain Sciences, 26(6), 682-683.