Crossing the divided line: Mathematical Perception and the Access Problem

Presented for the degree of Bachelor of Philosophy

Word Count: 28,887

By: Kevin Robert Riedel Gibbons



June 2018

Acknowledgements

There are many people to whom I owe a debt of gratitude for their help in the production of this thesis, or for helping me towards the production of this thesis. Unfortunately, they are too many to all be included in these acknowledgements, so I have to restrict myself to thanking those four people who have given the most of themselves.

First, I have to thank my supervisor, Alex, for his patience, insight, and kindness. He was willing to see the potential in me and my work, and went above and beyond the call of duty to see that potential realized. If I have grown as a philosopher during my time on the BPhil, it is only through Alex's guidance.

Next, I have to thank my Mom and Dad. As a kid, I was in the habit of thanking God for "every good thing" before meals because I knew I would not be able to remember every individual good thing I had to be thankful for (and also because I didn't want to put off eating while I tried!). I'm going to have to do something similar here, because I can't even begin to list all the things my parents have done for me. How do you measure the impact made by the people who brought you into the world, who raised you, who taught you to speak, and read, and think? No— it's impossible. All I can say is: thank you for everything.

Finally, I have to thank Brin, my wife and my best friend. Throughout my course, she was always there to support me when I needed to work, and she was always there to distract me when I needed a break. She also always reminds me that "there are more things in heaven and earth... than are dreamt of in your philosophy", which has helped me preserve (some of) my sanity. Thank you for always being up for an adventure, and for so much love and help along the way. I wouldn't have been able to do it without you.

Contents

1 The Access Problem in mathematical epistemology **1**

1.1 The Access Problem in the contemporary debate 6

1.1.1 Benacerraf's formulation 6

1.1.2 Field's generalization 8

1.1.3 Strong Gödelian response 18

2 The Number-Property approach to mathematical perception **20**

2.1 Maddy's Insights 20

2.1.1 One object, many views 21

2.1.2 Composition and Causation 24

2.1.3 From perception to intuition 27

2.2 A Problem: Are number-properties enough?? 29

2.2.1 Mathematical objects as spatial 33

3 A Structuralist Version of Weak Gödelian epistemology **36**

3.1 Structuralist epistemology 36

3.1.1 Structuralism and pattern recognition 39

3.1.2 From simple structures to complex 44
3.2 Cognitive basis of a Weak Gödelian argument 51
3.2.1 The psychophysical basis of ordinary perception 51
3.2.2 Philosophy from physiology 55
3.2.3 Spatiality: An objection 66

4 Information and Interaction **72**
4.1 Hart's Energy Argument and Maxwell's demon 74
4.1.1 Refining Hart's Condition 78
4.2 Physical information 80
4.2.1 Mutual information: Hart's Condition Revisited 84
4.3 A Weak Gödelian reply 86
4.4 Conclusions 96

Abstract

Platonism in the philosophy of mathematics holds that there are mind-independent, abstract objects which are the subject of mathematical inquiry. This view faces a vexed epistemological problem: given that humans are concrete, physical creatures squarely situated in space-time, how can we come to possess knowledge of these abstract objects? Gödel is famous for suggesting that we can have "something like a perception" of some mathematical objects, and this faculty for perception is how we gain (at least some) mathematical knowledge. The positing of such a capacity for mathematical perception was met with a critical reception. However, some contemporary philosophers have attempted to develop this project, including Penelope Maddy and Jaegwon Kim, as well as Max Jones more recently. These authors have gone a good way towards making mathematical perception more scientifically plausible. However, they have done so while working with a 'number-property' view— the view which says that '2' refers to a property instantiated by any two objects taken together. I argue that, for various reasons, such an account is undesirable. In this thesis, I investigate whether the project of basing mathematical epistemology on 'something like perception' can be divorced from the number-property view. I conclude that it can, and demonstrate why this is. The structure of the thesis runs as follows: In Chapter 1, I introduce the access problem and the perceptual approach to answering it. In Chapter 2, I look at the number-property approach to mathematical perception, and argue that it is wanting in a few crucial respects. In Chapter 3, I look at current research in cognitive neuroscience, and find it consistent with a number-object account of mathematical

perception. In Chapter 4, I formalize the access problem for the number-object view in terms of information theory, and find that it is less serious than has been previously made out.

Chapter 1

The Access Problem in mathematical epistemology

In Book VI of the *Republic*, Socrates introduces the metaphor of the Divided Line. The metaphor runs as follows: Consider a line, which represents all the objects there are. We can divide this line into two sections; this division represents the division of all the objects there are into two kinds. There are the visible objects, and there are the intelligible objects. The visible objects are those which are grasped by vision(or by any of the senses, generally speaking). The intelligible objects on the other hand are those which cannot be grasped by vision, but only by the understanding.

These two sections are further subdivided. The visible section is divided into the Originals and the Images. The Originals are things like 'the animals around us, all the plants, and the whole class of manufactured things' (*Republic VI*, 510a). The Images are things like 'shadows, reflections in water and in all close-packed, smooth, and shiny

materials' (*Republic VI*, 510a). On the other hand, the intelligible section is divided into the Mathematicals and the Forms. Examples of the Mathematicals include 'the odd and the even, the various figures, the three kinds of angles, and other things akin to these' (*Republic VI*, 510c). The Forms are things like Justice Itself, Beauty Itself, The Good, etc.

Famously, introducing such a distinction gives rise to epistemological problems. Parmenides warns a young Socrates of this in his eponymous dialogue, saying "whoever hears about [the Intelligibles] is doubtful and objects that they do not exist, and that, even if they do, they must by strict necessity be unknowable to human nature' (Parmenides, 135a). The question thus becomes: if there is another kind of object beyond those which inhabit the everyday world around us, how do we, flesh and blood human beings, come to know them?

This is the central epistemological problem for any metaphysics which posits the existence of abstract objects, and it has proven to be a difficult one. But to the contemporary philosopher with a naturalistic bent, such a problem may seem to be of marginal interest. After all, there is enough philosophical work to be done in interpreting the advances made by modern science and assimilating them into a coherent worldview. Why should we devote any time to obscure questions about how we come to know the Forms, when the only place you even hear about such Forms is in the philosophy classroom?

Unfortunately though, the problem is not so easily written off. Not only do the Intelligibles include the Forms, they also include the Mathematicals. While the naturalist may be happy to not say anything about the Forms, the significant role which the Mathematicals play in our scientific enterprise demands explanation. Given that the Mathematicals are abstract in the same way that the Forms are, providing such an explanation promises to

be challenging.

Still, I attempt such an explanation in this thesis. Or, to be slightly more accurate, I give an outline of such an explanation as a 'proof of concept'; in doing so, I hope to show that it is at least initially feasible. The explanation runs as follows: People can perceive basic mathematical structures. These perceptions yield knowledge of basic mathematical statements. From this body of perceptual knowledge, we come to know the axioms in a way analogous to how we come to know physical laws after observation. Knowledge of the axioms is therefore empirical on this picture. From the axioms, we proceed to gain knowledge of mathematical theorems via deduction in the standard way. The reader may feel that such an explanation sounds rather radical, but this impression is just a side effect of the generality with which we've sketched the position. As the details are filled in, we will find that the position is relatively sober.

It will be noted that in proposing such an account, I follow Gödel in positing "something like a perception" of some mathematical objects (Gödel, 1964). In doing so, I potentially open myself up to the kinds of criticisms Gödel originally faced. It will therefore be useful to make a couple of distinctions at this point. Anyone who follows Gödel and postulates a faculty of mathematical perception has one of two options. First, they might postulate a new, independent faculty of mathematical perception that is not related to our other perceptual faculties. Let's call this view the *Strong Gödelian* position. This position is often caricatured as claiming that the mathematician has some mystical power with which they access Platonic heaven and thus come to know about the mathematical objects which reside there. In contrast to this view, we have the *Weak Gödelian* approach. Rather than positing some new perceptual or intuitive power, the Weak Gödelian approach seeks

to explain mathematical perception by means of our normal perceptual capacities. For example, Penelope Maddy's argument in (1990) that we can perceive impure sets takes this line. Rather than positing some new power for mathematical perception, Maddy argues that the processes which give rise to ordinary instances of perception could also be seen as giving rise to mathematical perception. Thus, we assimilate mathematical perception into a previously established scientific framework, a hallmark of the Weak Gödelian approach.

The above are two distinct accounts one can give of the *faculty* of mathematical perception, but there is the further question about the *objects* of this faculty. In other words, what is the ontological status of the objects we come to know about via mathematical perception? Again, there are a number of routes to take. One route would be to maintain the standard assumption that mathematical objects are abstract objects which do not enter into any relations to objects in spacetime. On the other end of the spectrum, we could postulate that the objects of mathematical perceptions are concrete, and have a definite position in space-time (cf. Maddy, 1990, p.59). Additionally, one could pursue a hybrid route where mathematical objects are neither purely concrete, nor purely abstract. To make meanings explicit: I understand an object to be abstract iff it is non-spatial and causally inefficacious.

The position I want to defend is a Weak Gödelian account of mathematical perception, with a hybrid ontology of mathematical objects. I hold, in line with Penelope Maddy's previous writings (Maddy 1980, 1990), that we can understand our normal perceptual faculties as yielding perceptual knowledge of basic mathematical objects. Or, more specifically, I hold that any assumption which is strong enough to rule out mathematical perception (in basic cases) will also rule out cases of ordinary perception. In as far as the latter is unde-

sirable, we are compelled to accept the former. I make the parallel argument with respect to ontology: recognizing that ordinary objects already are surprisingly theoretical entities, I propose that basic mathematical objects differ from them only in degree and not in kind. Thus, they are just as eligible to be the objects of perception as ordinary objects are.

I am aware that these two proposals may seem to be implausible, or indeed, baffling from a scientific point of view. Much of the literature on mathematical knowledge registers a deep confusion about the sort of interactions which could possibly obtain between human knowers and mathematical objects. We often hear derisive mention of 'signals' coming from the mathematical realm (MacBride, 2008) or 'energy transfer' happening between humans and mathemtical objects (Hart, 1977). On closer inspection though, such reasoning amounts to little more than rhetorical posturing. In fact, when we review the relevant sciences we see that they are not only compatible with mathematical perception, but in some cases even seem to suggest it. Following on work done in Maddy (1980, 1990) as well as Jones (2015), a great deal of the present thesis is concerned with demonstrating this compatibility of a Weak Gödelian account with the relevant scientific theories.

Though the details have yet to be filled in, the general dialectical structure should be relatively clear by now. In response to the question 'How do we come to possess mathematical knowledge?', I answer 'in part, by perceiving mathematical objects'. In giving this answer, I emphasize two things: 1) If ordinary objects can be perceived, basic mathematical objects can be perceived as well 2) Mathematical objects do not differ in kind from ordinary objects in anyway which allows us to draw a principled distinction between the two with respect to perception. To put it more succinctly, I argue that the middle division in Plato's divided line is a false one; The Orignials and at least some

Mathematicals do not differ in what kind of thing they are, nor in how they are known.

The rest of the thesis is structured as follows: In the remainder of Chapter 1, I present the access problem in greater detail, and in its contemporary form. After rejecting the strong Gödelian response, we proceed to Chapter 2, in which I review the Weak Gödelian literature up until this point. While I find much to agree with in this literature, I conclude that current Weak Gödelian theories fail to satisfy certain desiderata for an adequate philosophy of mathematics. In particular, I hold that the number-property account of mathematical perception fails to account for the diversity of mathematical knowledge. In response to this, I develop the Weak Gödelian project within a structuralist framework in Chapter 3. Chapter 4 defends the idea that the faculty of mathematical perception described can access mathematical objects, using insights taken from information theory.

1.1 The Access Problem in the contemporary debate

Though the access problem in mathematical epistemology has ancient origins, its modern formulation is found in Paul Benacerraf's 1973 paper *Mathematical Truth*. But canonical formulations rarely last long in philosophy, and Benacerraf's is no different. So after giving Benacerraf's statement of the access problem, we will look at some more recent literature on the problem. Doing so will yield a precise formulation of the problem we wish to address in later sections.

1.1.1 Benacerraf's formulation

Consider two sentences:

1. There are at least three cities larger than New York.

2. There are at least three natural numbers greater than 17.

On the face of it, both of these sentences have the same grammatical form.[1] Specifically, they are both of the form 'There are at least x F's that bear R to k'. On a Tarskian account of truth and satisfaction then, (1) will be true iff there are three or more objects of the domain which satisfy the 'city' predicate and which stand in the 'is larger than' relation to the object denoted by 'New York'. If the model is taken to be the real world, and the interpretation is the correct one for English, these satisfaction requirements are met, and therefore (1) is true.

As Benacerraf notes, there is a strong urge to give a similar account of the truth conditions of (2). After all, if we are going to give a homogeneous truth theory for sentences in our language, then it is desirable to give parallel truth conditions for sentences with parallel grammatical structure. Lacking such homogeneity, we would not only need to provide a truth theory, but also a truth theory theory which would explain why we give different truth conditions for sentences which are similarly structured. So concerns about consistency of explanation may lead us to analyze (2) in the same way as (1).

This way of spelling out the truth conditions for mathematical sentences seems to push us towards a Platonist approach to mathematical objects, assuming we take (2) to be true. To see this, assume (2) is true. Thus, there are at least three objects in the domain which satisfy the predicate 'is a natural number' and which bear the 'is greater than' relation

[1] An opponent of the 'standard' approach to be sketched here might respond by drawing a distinction between the surface grammar of these sentences, and the real grammar of these sentences (cf. Deep Structure and Surface Structure in Chomsky(1964)). I won't consider this objection in much detail here, but flag it still.

to the denotation of '17'. Which objects exactly are these? Which objects are in the extension of 'is a natural number'? What is the denotation of '17'? To answer these questions in a sensible way, we are forced to posit the existence of 17 and the rest of the natural numbers too. This is to adopt a Platonist ontology of mathematics. Benacerraf refers to this mixture of a Tarskian theory of truth with a Platonist ontology the *standard account* of mathematical truth.

After outlining and showing the advantage of the standard account of mathematical truth, Benacerraf considers how compatible it is with a satisfactory account of mathematical knowledge. He is not optimistic, writing:

> The principal defect of the standard account is that it appears to violate the requirement the requirement that our account of mathematical truth be susceptible to integration into our overall account of knowledge. (1973, p. 670)

The problem, as he sees it, it this: for X to know sentence S, we need 'some causal relation to obtain between X and the referents of the names, predicates, and quantifiers of S'(1973, p. 671). But given that the standard account takes the referents of the terms in mathematical sentences to be causally inert, we seem completely barred from mathematical knowledge of any kind. Thus, we seem unable to give a satisfactory account of mathematical truth and mathematical knowledge together. This is Benacerraf's problem.

1.1.2 Field's generalization

It is generally held that Benacerraf's original formulation of the problem is too reliant on the causal theory of knowledge, and that this reliance causes problems. The main arguments against causal conditions on knowing in the context of mathematical epistemology are

found in Steiner (1975, Chapter 4). Steiner's point is this: the causal theory of knowledge needs specification. Consider a case where S believes that P. For S to know that P we could frame the causal condition as requiring: 1) the fact that P cause S's belief that P 2) the subject of P cause S's belief that P 3) an event in which the subject of P features cause the knowledge 4) that P feature in a causal explanation of S's belief that P. Steiner argues that causal arguments face a dilemma: either the causal condition endorsed is implausible, or the causal condition fails to rule out Platonism.

I am dubious of Steiner's argument for this dilemma. He takes the only plausible version of a causal condition to be the one specified in terms of explanation. However, he thinks that this is the causal condition which is compatible with Platonism, because any 'theory, actual or possible, known or unknown, which can satisfactorily explain our belief in causal style...WILL CONTAIN THE AXIOMS OF NUMBER THEORY AND OF ANALYSIS [sic]' (1975, p.114). Of course, whether Steiner is correct in this claim will depend on what counts as a satisfactory causal explanation. It appears as if Steiner has in mind an explanation in a scientific language which includes a good bit of mathematics, but he doesn't explicitly say. One could easily resist this classification, allowing descriptions in ordinary English (or even scientific jargon with no mathematics) to count as satisfactory causal explanations. If this is allowed though, then it isn't clear that the truth of a given bit of mathematics will feature into a causal explanation of some subjects belief in that bit of mathematics. Thus, the most plausible causal condition may be incompatible with Platonism after all.

Still, I think the anti-Platonist should concede the point, and not tie their objection up with the causal theory of knowledge. It would be better to give an argument against

Platonism which wasn't dependent on specific views about causation and ontology. Such a line is pursued by Field, who writes:

> [I]f one postulates mathematical entities, one is going to have a serious problem explaining the correlation between our mathematical beliefs and the facts about those entities; and I have tried to argue that *to the extent that providing such an explanation appears impossible,* believers in a mathematical realm face a genuine epistemological problem... however good the reasons for believing in mathematical entities, the difficulty in explaining the reliability of those beliefs remains, and it raises a serious epistemological puzzle for those not happy with taking a perception of the realm of mathematical entities as primitive. (1989, p. 238-239)

By framing the question in this way, Field shifts attention away from what causes mathematical knowledge to what justifies it. We can grant that our mathematical beliefs are correct, and we grant that they are formed reliably. We even grant, for the sake of argument, that Platonism about mathematical objects is true. But if Platonism *is* true, and mathematical facts express truths about abstract and causally inert objects, then how can we explain the reliability of our mathematical beliefs? In other words, how can we explain the following conditional:

Mathematical reliability claim. For any mathematical statement S, if mathematicians accept S, then S is true.[2]

If we can't explain this conditional on the assumption that Platonism is true, then

[2]I'm using talk of whether 'mathematicians accept S' as a loose approximation for what the majority of working mathematicians who specialize in the areas pertaining to S think about S.

that seems to count against it. Thus, Field's formulation of the access problem runs like a reductio argument. We assume for sake of argument that Platonism about mathematical objects is true. If that is the case, then providing an adequate explanation of the mathematical reliability claim (MRC) 'appears impossible' (ibid.). Since the MRC is true (we assume), Platonism precludes explanation of a true fact.

Moreover, it seems we can know that Platonism precludes explanation of a true fact. But consider this doxastic principle: Don't believe propositions which you know make it impossible to explain a true fact. If this principle is correct, we can conclude that we ought not believe in Platonism. So: is it correct? Well, first of all, we should clarify what meant by a proposition making an explanation of another proposition impossible. Suppose F is some field of study with a collection of foundational assumptions S. Then we say that P makes an F-explanation of Q if P, Q, and S are mutually inconsistent. Since sets of true sentences cannot be inconsistent, the assumption that S and Q are true implies the falsity of P. So, on this precisification, the doxastic principle to not believe propositions which makes explanation of true facts impossible follows directly from the principle to not believe false propositions. Applying this to the MRC, it seems we are compelled not to believe Platonism. Thus, Field's formulation jettisons the unnecessary assumptions found in Benacerraf, while retaining much of the argumentative power.

However, though Field's formulation is more streamlined, it does make two substantial assumptions. One is that Platonism precludes an explanation of the MRC. The other is the doxastic principle that we should not believe things which make explanations of true facts impossible. In Clarke-Doane (2016), we see that these two assumptions give the Platonist room to push back against Field's formulation of the access problem. The tack

the Platonist can take is to demand a precisification of the notion of reliability involved in Field's argument, and to argue that this demand cannot be met satisfactorily. Clarke-Doane writes:

> Nevertheless, [Field's formulation] is unclear at a crucial juncture. It is unclear what it would take to explain the reliability of our mathematical beliefs in the relevant sense. In what sense of 'explain the reliability of' is it plausible both that it appears in principle impossible (for us realists) to explain the reliability of our mathematical beliefs, and that the apparent in principle impossibility of explaining the reliability of those beliefs undermines them? (2016, p. 22)

If that is the case, then it seems that the access problem, on Field's interpretation at least, isn't specific enough to be a serious threat to Platonism. We now sketch this response from the Platonist, to see whether it is enough to put the access problem to bed.

Safety and Sensitivity

The first precisification of reliability we'll consider uses the notion of a belief's safety. We suppose we have explained the reliability of our mathematical beliefs if we have explained that they are safe. A safe belief is defined as a belief that could not have easily been wrong in a similar case. Of course, we immediately want to know what is meant by 'easily' and what counts as a 'similar case'. We specify this as follows: for any true belief P held by S in the actual world, we say that S's belief P could have easily been wrong if there is a nearby possible world in which S still believes P, but P is false.

If it is granted that mathematical truths are metaphysically necessary, then our true mathematical beliefs must meet the first criterion. After all, if there are no possible worlds

where P is false, then clearly there are no possible worlds where P is false and S believes P. Thus, on at least one precisification of reliability, it is rather trivial to explain the reliability of our mathematical beliefs.

In response, one could try to cash reliability out in terms of sensitivity. S's true belief in P is said to be sensitive if it varies counterfactually with the truth of P- if P hadn't been true, then S wouldn't have believed P. Since, again, mathematical facts are necessarily true if true at all, it seems at first that our true mathematical beliefs are trivially sensitive. Since there are no worlds where P is false, there are no worlds where P is false and we believe that P. Or at least, one might argue that way at least. Field, however, resists the idea that we can't make sense of the counterfactual behavior of necessary truths. To see this, consider the sentences (i) 'If the axiom of choice is false, then the moon is made of green cheese' and (ii) 'If the axiom of choice is false, then the Banach-Tarski theorem is true'. Assuming the truth of the axiom of choice, then there are no worlds in which either of these statements are false, given that the antecedent is necessarily false. So in this regard, both of these sentences are trivially true. However, (ii) seems to also be wrong in an interesting way. Since the axiom of choice is required for the Banach-Tarski result to obtain, it seems like we have described a doubly impossible scenario. Not only have we violated the mathematical necessity involved in the truth of the axiom of choice, but we've violated the logical necessity linking the axiom to the Banach-Tarski theorem. On the other hand, I don't know of any proof which shows, given the negation of the axiom of choice, that the moon is made of green cheese, so no additional logical necessity has been violated in (i).

Thus, Field concludes, in some cases we can reasonably discuss some counterpossibles.

It remains to be seen whether the counterpossible (iii) 'If 2+2 didn't equal 4, then we wouldn't have believed that it did' is one such counterpossible. But, lacking an argument that it isn't, we are not yet entitled to conclude that our mathematical beliefs are sensitive.

Clarke-Doane does have an argument to this effect though. He grants that (iii) is like (ii) in that we can meaningfully discuss about it being wrong. But he then says something interesting, which I think is very much along the right lines, writing:

> We seem to be equally unable to show that our relevant uncontroversial beliefs are non-vacuously sensitive. For example, we seem to be unable to show that our belief in 'bridge laws' that link supervenient properties to subvenient ones are so sensitive. Had- as a matter of metaphysical impossibility- atoms arranged car-wise failed to compose cars (as "ontological nihilists" allege), we still would have believed that they did. (2016, p.26)

Thus, if we deny the sensitivity of beliefs with metaphysically necessary antecedents, we have to deny the sensitivity of important beliefs about the composition conditions of ordinary objects. If our beliefs in these composition conditions are undermined, then this also undermines our beliefs about the objects themselves. As he says, 'If our belief about the composition conditions of cars is undermined, then so too, presumably, is our belief that we are sitting in one' (ibid). In as far as we don't want to undermine our beliefs about ordinary objects, we cannot consistently reject the sensitivity of mathematical beliefs and accept the sensitivity of ordinary beliefs. Thus, the access problem fails to be unique to mathematical knowledge.

This is a very important point which deserves emphasis. The correct formulation of

the access problem cannot generalize to ordinary objects. If it does so, then it introduces a global skepticism and becomes philosophically untenable. Echoing a point made in Steiner (1975), we must accept that, if we are to avoid undue skepticism, *we cannot reject the existence of mathematical objects for reasons which would also lead us to reject the existence of ordinary objects.* This is a theme which will be picked up and emphasized throughout this thesis. Clark-Doane does well to bring this fact to the fore, and to relate it to the recognition that the existence of ordinary objects requires certain metaphysical relationships to obtain (i.e. composition). We will return to composition relations and their role in the mathematical epistemology debate in Chapter 4, but for the time being we can grant that the access problem is not obviously about sensitivity.

The Access Problem and Explanation

So sensitivity goes the way of safety, and we still don't have an adequate precisification of the 'explain the reliability' requirement posed by Field's formulation. Rather than conclude that this failure to formulate the access problem in terms of sensitivity or safety exposes it to be a pseudoproblem, we follow the more appropriate approach found in Linnebo (2006). In this paper, Linnebo outlines an improved epistemological question. The new challenge to the Platonist isn't to simply show that mathematical beliefs are reliable (which Linnebo would call an internal explanation), but rather for them to explain in a scientific manner what makes it the case that they are reliable. As Linnebo writes:

> Just as it makes perfectly naturalistic sense to require more than an internal explanation of perception, the same requirement makes sense of mathematics as well. This requirement is neither contrary to science nor outside its scope;

> rather, it flows from science itself. We want an account of how the central epistemological notions, such as evidence, reliability, and justification, fit into a scientific account of the epistemic agent and his natural surroundings. (2006, p.566)

The challenge to the Platonist isn't to show that mathematical beliefs are safe, or sensitive, or reliable in some other way. Rather, the challenge to the Platonist is to show that a scientific explanation of mathematical knowledge is possible. The access problem arises from the fact that there is a discrepancy in language: a scientific explanation would, presumably, use the vocabulary of physics, biology, psychology, etc. But how can this vocabulary make any headway in describing how we come to know about abstract objects? We could state the improved access problem as follows: If Platonism makes a scientific explanation of mathematical knowledge impossible, it should be rejected. Since it does, it should be.

One naïve objection to this requirement runs as follows: This formulation of the access problem confuses psychology with epistemology. It is the job of psychology to give an explanation of the processes by which we form our beliefs about certain objects. Further, it should do so in a way that is congruous with the other natural sciences. But epistemology does a different job. Its objects of study aren't the natural processes which may or may not give rise to knowledge, but rather knowledge itself. But knowledge, properly speaking, is not a scientific concept; rather, it is a philosophical one. Since it is not a solely scientific concept, why should we think that our explanation of knowledge, and mathematical knowledge in particular, needs to be congruous with the natural sciences? If there is no good reason for thinking so, then we should be free to give defenses of mathematical knowledge

like the ones Clarke-Doane gives: we show that mathematical knowledge is *philosophically* unproblematic, and that is all that needs to be said about the matter.

The response one might give to such an argument goes like this: first we try to counter the idea that knowledge is strictly a philosophical, rather than a scientific, concept. We grant that epistemology will need to utilize extra-scientific notions (i.e. notions which go beyond those found in the natural sciences), but we can still deny that all of the notions in epistemology need to be extra-scientific. There is good reason to do so. Most importantly, epistemology requires that there be subjects which possess knowledge. Usually, these subjects are taken to be relatively normal human beings; in other words, we are the subjects. But there are obviously many other fields which study, at least partially, human beings. Biochemistry, anatomy, anthropology, neurology, and a host of other fields in the natural sciences tell us interesting things about ourselves. All of these fields will have unique notions and methods, but we don't take this to mean that these fields don't need to be somehow compatible with one another. Rather, we take it to mean that they tell us about different aspects of the same human beings. For the epistemologist to try to divorce his field from the natural sciences does more harm than good then: it may obviate the need to provide an external explanation of the kind Linnebo demands, but it also robs the field of much of its interest. It is as Pope writes: 'The proper study of mankind is man'. In as far as mankind is an object of scientific explanation, epistemology must be congruous with the sciences, or fail to be a 'proper study'.

1.1.3 Strong Gödelian response

Still, it must be granted that formulating the issue in this way yields a compelling version of the access problem: We cannot be Platonists about mathematical objects and give a plausible scientific explanation of how we come to know about these objects. What can we say to this?

The line of response we are concerned with here is the Gödelian one. As mentioned above, there are two kinds of Gödelian views one can hold. The Strong Gödelian view posits a mathematical perception (or perhaps an intuition) which is distinct from our normal faculties of perception. Gödel seems to have held this stronger stance, as indicated by the following passage in (1964):

> [D]espite their remoteness from sense experience, we do have something like a perception also of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true... It should be noted that mathematical intuition [i.e. the faculty 'like a perception' which we have] need not be conceived of as a faculty giving an immediate 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 is not, or not primarily, the sensations. (1964, p.268)

There are many things that are unclear about Gödel's proposed faculty of mathematical intuition. For example, Burgess (2014) argues that there are actually three mutually inconsistent versions of mathematical intuition at play in Gödel's writings on the continuum

hypothesis. But even if it were possible to construct a unified account from Gödel's own writings, or to present an account independent from that of Gödel (but along the same lines) it is not clear how such an account would fare in the face of Linnebo's improved epistemological challenge. As Linnebo notes, it is fairly straightforward to give an external explanation of ordinary perception. Objects in our environment interact with various particles, then our senses are impressed by these particles, and these impressions are relayed to the brain, which then interprets the messages it receives to determine which objects are actually in the environment and what actions to take in light of this.[3] But what is the parallel with mathematical intuition? Gödel has explicitly said that the source of the 'mathematical given' is not sensation. What is it then? If there is no scientifically respectable source for the 'mathematical given', then an external explanation of the MRC seems impossible on Gödel's view. I can see no such source though, and so we tentatively conclude that Gödel's view cannot meet Linnebo's new epistemological challenge. Indeed, it seems right to conclude that any Strong Gödelian account which separates the faculty of mathematical perception from sensation will likewise be unable to give an external explanation of the MRC. Therefore, if we wish to appeal to mathematical perception to solve the access problem, we should give an account which posits mathematical perception as a further aspect of ordinary perceptual systems which admit of external explanation. In the next chapter, we review existing proposals of this kind.

[3]There are obviously questions of how the things like consciousness supervene on these physical processes, but this doesn't undermine the validity of asking for a physical explanation

Chapter 2

The Number-Property approach to mathematical perception

2.1 Maddy's Insights

One cannot give a good philosophical treatment of mathematical perception without discussing Penelope Maddy's work on the subject. In a series of papers published in the 1980's, and then especially in her 1990 book *Realism in Mathematics*, Maddy gives the seminal presentation of the Weak Gödelian programme. This presentation contains three important insights which will be crucial for our discussion going forward. They are as follows: first, object perception requires the the identification of many different sensations as *of the same object*. Just as important as this first insight is the argument used to support it, which we will examine shortly. Second, in any given case of perception, the sense organs can only be said to causally interact with 'part' of the object. Lastly, the

same neurophysical processes which lead to perceptual beliefs can also result in intuitive beliefs about general concepts. This allows her to explain our knowledge of the axioms of set-theory by appealing to perception. Each of these insights is essential for defending the Weak Gödelian approach, and so it will be worth looking at them all more closely.

2.1.1 One object, many views

Consider smoke detectors. Smoke detectors are made with one end in mind: alerting their users to the presence of a substantial amount of smoke in the air around them. To do this job, a smoke detector must be able to carry out two functions. First, it must be able to recognize when the smoke in the air has exceeded a certain threshold. Second, it must be able to sound an alarm when this criterion has been met.

Neurons are like smoke detectors in that they perform the same twofold function of detecting and alerting. But in the case of neurons, what is detected isn't the presence of smoke in the air, but rather the presence of certain chemicals in the surrounding intracellular fluid.[1] And instead of alerting by emitting loud, high pitched beeping noises, neurons alert by emitting chemicals of their own into the intracellular fluid. In this way, what neurons detect is also very similar to what they emit. Because of this, neurons can not only be thought of as detectors of intracellular chemicals, they can also be thought of as detectors *of activity in other neurons.*

In fact, not only are neurons well suited to detecting activity in other neurons, they are well suited to detecting particular patterns of activity in groups of other neurons. So, for example, let's consider an electronic 16x16 grid, where each cell on the grid can be

[1] Or, for specialized nerves cells in the retina, the presence of photons.

illuminated or dark. Let's assume, to simplify things, that a system of neurons responds perfectly to this illumination. So there are 256 neurons in this system, and each one fires exactly when its corresponding grid cell is illuminated. Because neurons are also good at detecting particular patterns of activity among other neurons, we can consider another system of neurons in addition to the first. Let's suppose these neurons are configured to detect specific patterns of activation in the first layer. In particular, let's suppose that they are each attuned to fire when the activation pattern in the first layer corresponds to the grid displaying one of the digits 0-9. Thus, a neuron's ability to detect activation patterns[2] in other groups of neurons allows us to construct different layers of detection. The type of 'thing' detected by a particular neuron will depend on which layer it detects activations in.

The above example is highly simplified, but in essence it is the same kind of process which is responsible for vision. The main differences are that the 'grid' is much larger in the case of vision, and there are more layers of abstraction. We will go into this in more detail in Chapter 3, but for now it is sufficient to think of vision according to this grid analogy. It is tempting to think that at this stage we have all that is needed for visual perception. After all, we have neurons which respond to sensory input, and a layer of neurons which is responsible for detecting different figures in this sensory input. What more do we need?

By itself though, this neurophysical story isn't rich enough to account for perception. What is needed is more than just a layer of neurons which detects figures in our sensory input. To have a bona fide perception of an object, we need to be able to recognize the different figures which this object can bring about in our sensory experience. By way of example, consider my favorite chair. If I look at this chair in a well lit environment, I will

[2]If detect an activation pattern is a little bit of a loose way to speak, we could be more precise by saying that neurons can detect the activation of definite groups of other neurons.

be able to discern its figure. Now suppose I walk to the other end of the room, and look around at my chair again. Strictly speaking, the figure which presents itself to my sensory experience will not be the same as before. At the very least, the chair-figure will be smaller than at first— it will occupy less of my visual 'grid'. To recognize that these two figures given in my sense experience correspond to the same object, there needs to be another neuronal layer on top of the figure-detecting one. This layer will need to detect when any figure corresponding to my chair is within the visual field. In short, if I am to going to perceive my chair at all, then there needs to be a 'this chair' neuron (or cluster of neurons) which ignores irrelevant features like distance, angle, etc. Thus, to perceive an object at all, we need to be able to identify all the different sense experiences of that object as the same *qua* object.

So that is where the physical basis of object perception lies: a few layers up in this sequence of neuronal processing. It should be clear by this point that this neurophysical mechanism is particularly good at one thing: transforming manifold inputs into a unity. At the first stage, manifold sensations are unified as one figure. At the second stage, manifold figures are unified as one object. Maddy's single most important insight is that this sequence doesn't need to stop at the level of ordinary objects. She writes:

> Given the evidence that the set concept requires a similar developmental period involving repeated experiences with sets in the environment parallel to the required experiences with... physical objects, it seems reasonable to assume that these interactions with sets of physical objects bring about structural changes in the brain by some complex process...the resulting neural 'set-detector' is what enables adults to acquire perceptual beliefs about sets. (1990, p.65)

The complex process Maddy refers to is the process which conditions neurons to detect certain activations in other neural layers. The picture which Maddy wants to paint runs as follows: we have some physical objects, say three eggs, and a subject (Steve) observing them. These objects impact Steve's sense organs via some intermediate physical process. In doing so, they stimulate a certain array of retinal nerves to fire. This firing is detected by certain neurons in the visual cortex, which are responsible for processing figures, edges, shading, etc. The firing of these 'figure' neurons is in turn detected by neurons in a higher layer, which are responsible for object identification. Maddy suggestion is that there is another layer on top of this one, which detects certain activations among the object detectors. Thus, when three of the egg-detector neurons go off, a neuron in this last layer may be said to detect the presence of an egg set.

This, then, is the point: The same neurophysical process is responsible for our coming to associate many figures with one object, and for our coming to associate many objects with one (impure) set. If the former can act as the physical basis for perception, we should accept that the latter can do so as well.

2.1.2 Composition and Causation

So the first insight Maddy had is that the neurophysical structure of perception is *prima facie* compatible with perception of impure sets. Her second insight[3] lies in her defense of the first. She considers the following objection: it's not anything about the perceptual process which makes perception of sets impossible. Rather, it is the nature of sets that makes perception of them impossible. Specifically, that they are causally inert. Though

[3]Well, in (1990) she attributes the insight to John Burgess, but the credit goes to her for developing it.

the causal theory of knowledge may not be plausible, a causal theory of perception is. After all, how can we be said to perceive something if it doesn't interact with any of our sense organs? But our sense organs, being physical, can only have interactions with other physical objects. Sets, whether pure or impure, are non-physical though, so they cannot interact with our sense organs in the appropriate way. We therefore conclude that we cannot perceive sets.

Maddy makes two moves in response to this argument. The first is to point out that this criterion is far too strong, since it would rule out our having perceptions of ordinary objects as well. After all, we don't causally interact with objects, *per se*. As Maddy writes:

> The interaction is actually with the front surface of a time slice of the sample... Similarly, the set-dubber has only aspects of her sample sets withing her causal grasp. But, if the interaction of the gold-dubber with an aspect of her samples is enough to allow her to pick out a sample and dub a kind, why shouldn't the set-dubber's interaction with an aspect of her samples accomplish those same things. The Platonist could argue that the relation of the element to set is no more objectionable than the relation of fleeting aspect to temporally extended object. (1990, p.49)

Since we never actually causally interact with the whole object of our perception, only with the slice of it which directly faces us, then we would have to rule out cases of ordinary perception. However, if we allow that we can perceive objects by causally interacting with their 'aspects', as Maddy says, then we can also interact with sets by interacting with their members.

We should note the ambiguity in the term 'aspect'. In the physical case, 'aspect' seems to mean part. We don't interact with the whole object, we interact with that part (or those parts of it) which are immediately visible to us. For most three-dimensional objects, this part will be a proper part of the whole, as we've noted. But in the case of perceiving sets, the idea that we even interact with parts of that set isn't immediately obvious. For example, we might follow David Lewis' arguments in *Parts of Classes* (1991) and hold that 'one class is a part of another iff the first is a subclass of the second' (p.4). On this view though, the analogy between sets and physical objects doesn't seem to help. The fact that we only need to interact with a part of an object to perceive it doesn't help us explain our perception of sets if the parts of a set are themselves sets. Such an explanation would be circular.

So Maddy denies this view. Instead, she holds that the parts of a set are its members.[4] Assuming that parthood is transitive, when we interact with part of the egg, we thereby interact with part of the set of three eggs. But note that this leaves Maddy's account of sets in a strange position. If a set has parts which are spatiotemporal, then the set itself seems to be a spatiotemporal entity. In fact, this is a conclusion Maddy is willing to make, writing:

> A second mode of resistance is to join the traditional Platonist in denying that sets have location in space or time. But notice: there is no real obstacle to the position that the set of eggs comes into and goes out of existence when they do, and that, spatially as well as temporally, it is located exactly where they

[4] Actually, the term Maddy uses is 'aspect', but it is not exactly clear to me what an aspect is. I proceed taking it to mean something like part, but flag that this could be an incorrect interpretation.

are. (1990, p.59)

So Maddy is willing to admit that impure sets at least are located in spacetime, in exactly those regions which are occupied by it members. I think this is a downside to Maddy's version of Weak Gödelian epistemology, as it has some negative consequences which we'll look at in §2.2.1. In fact, the main motivation for the alternative account to be developed in Chapters 3 and 4 is separating the Weak Gödelian project from the idea that the mathematical objects perceived are spatiotemporal. Even in this alternative treatment though, the recognition that the thing (or things) sensed doesn't need to be identical to the thing perceived plays a crucial role. We therefore still need to emphasize this as an important insight on Maddy's part.

2.1.3 From perception to intuition

So Maddy's defense of set-perception has motivation, but would appear to come at the cost of spatiality. Maddy's third and final insight shows that the benefit to set-theoretic realism may outweigh this costs. The reason for this is that set-theoretic realism provides an account of how we come to know the axioms of set theory. If this account works, then the rest of mathematical knowledge is supposedly secured, so long as it follows from these axioms.

What does the account look like? Recall the neurophysical picture: We have our first layer of neurons responsive to sensory input (the sense layer), then a second layer responsive to excitation patterns in the first layer (the figure layer), and then a third layer responsive to excitation patterns in the second layer (the object layer). But what's to say that this sequence stops there? Maddy quotes Hebb, who suggests that there is evidence for a

fourth layer of cell assemblies. Maddy explains that '[t]his fourth-order assembly would correspond to the general concept of a physical object. It would be stimulated during the phase sequence associated with perception of a particular physical object when attention is drawn to its more general features' (1990, p.69). Just as the development of an object concept allows us to come to know things about that object, the development of this fourth layer then which allows us to learn general facts about physical objects. For example, I must possess the concept 'my dog Spot' before I can come to know things like 'Spot is loyal' or 'Spot stinks'. Otherwise, I wouldn't be able to tell that this loyal and stinky figure which keeps popping up in my sense experience is in fact the same dog, Spot. Similarly, I need to possess the concept of physical object before I can come to know general facts about physical objects. If possessing a group of neurons which detect a thing is what amounts to possessing a concept for that thing, then having a group of neurons which responds when any physical object is perceived constitutes having a concept of physical object.

To carry on with Maddy's general theme, as go physical objects, so go sets. She posits the existence of a higher-order cell assembly which is responsive to the firing of set-detecting assemblies. Just as the physical object assembly yields knowledge of general facts about the nature of physical objects, this assembly yields knowledge of general facts about the nature of sets e.g. that any two sets can be combined to make another set, with the first two as members. These general beliefs, derived from our experience of impure sets, gain a more privileged epistemological status through empirical confirmation. Eventually, we lay them down as axioms which we assume to hold of all sets, pure and impure. From these axioms, we can come to know a great deal more about sets (and, in as far as other branches of mathematics reduce to set theory, about the rest of math as well). This additional

knowledge is, for the most part, unproblematic, given that it proceeds deductively.

2.2 A Problem: Are number-properties enough??

So these are Maddy's insights: that the neurophysical nature of perception is compatible with perception of sets, that sets do not appear to be more causally remote from sensation than ordinary objects of perception are, and that perceptual knowledge of particular sets can give rise to general knowledge of sets, in general, which forms the basis for the axioms of set theory. Like I've said, there is a lot to like in this account. One might be tempted, at this point, to conclude that the job is now done. After all, most of mathematics can be reduced to set theory. If we have given a viable epistemological account of set theory, we can be said to have shown how mathematical knowledge is possible.

To rest content at this stage would be a mistake though. We have only shown that there is a viable route for us to gain knowledge of those mathematical facts which we know. But the adequacy requirements for an epistemology of mathematics are starker than this; not only do we need to show that it is possible to know those truths which we actually know, we need to show that it is possible to know those truths which we actually know *via the methods which we actually know them.* In light of this requirement, the story told up to this point fails. People can and do have arithmetical knowledge without having to derive it from general truths about sets. For an epistemology to be complete, we must be able to explain this knowledge as well.

Maddy is aware of this requirement. She writes:

> Assuming that numbers aren't sets, the set theoretic realist faces the prospect

> of adding a new type of entity to her ontology an extra epicycle to her epistemology. And, to preserve the explanatory force of the standard set theoretic reduction, she must also describe a relationship between sets and numbers that makes the behaviour of, for example, the von Neumann ordinals somehow relevant to our understanding of numbers. Finally, the efforts to naturalize the epistemology of set theory will be wasted if the account of numbers isn't also naturalistic. (1990, p. 87)

To meet this challenge, Maddy proposes that '[k]nowledge of numbers is knowledge of sets, because numbers are properties of sets' (ibid., p.89). So for example, the impure set containing the three eggs in Steve's refrigerator has the property of *being three*. Just as we can come to know about normal properties by perceiving the objects which instantiate them, so too can we learn about numbers by perceiving the sets which instantiate them. Or at least, that is how the view is supposed to run. In actuality, the move to viewing numbers as properties causes problems.

The first problem runs like this. The primary motivation for Platonism about mathematical objects is semantic: we want to be able to assign singular terms in our mathematical languages actual objects as their semantic values. But, if the object vs. concept distinction is maintained, and numbers are relegated to the concept side of the distinction, then the semantic motivation for Platonism is undercut. We cannot give a semantics for mathematical language which respects the surface grammar of that language.

Maddy is unmoved by this issue though, writing 'grammar is no infallible guide to the actual structure of the world, so such evidence must be taken with a grain of salt' (ibid., p. 90). If the set-theoretic realist's account is otherwise attractive, then we should be

willing to accommodate some paraphrasing. But I think this sort of move obscures the actual issue at hand, which relates to what we said earlier. If the grammar of mathematical language is correlated with the nature of mathematical practice, then we will have to respect grammatical structure if we want to give an accurate account of how we actually come to possess mathematical knowledge. It might be that consideration of sets and their properties doesn't enter into our arithmetical judgements at all. In that case, an epistemology which relies on the assumption that numbers are properties of sets cannot but fail to account for how we come to make true arithmetical judgements. If that's so, then we have failed to meet the criterion which we posited above; namely, that our account of mathematical knowledge explain how actual people come to possess the mathematical knowledge that they do.

But let's grant for a moment the assumption that some leniency can be accepted when interpreting the grammar of mathematical language. How useful is the proposal that numbers are just properties of sets? On closer inspection, it is not clear that it very useful at all. First of all, given that all the sets which we perceive are impure, we will not come to the notion that there are empty sets. If numbers are only properties of sets then, there will be no zero.

And what about the rest of mathematics? What about the mathematical objects with more than one relation? How are we to account for them? After all, set-theory only has one binary relation. But many mathematical objects require more than one binary relation to adequately describe. For example, the ring of integers has two binary operations defined on it, addition and multiplication. Where does our knowledge of these operations come from, given that the objects which they operate on are actually properties? In short, it seems

that sets are only rich enough in structure to interpret other mathematical theories, they are not rich enough to suggest them of their own accord. But, if numbers are properties of sets, then that's what we would expect.

It may be objected that this criticism underestimates how rich a set's structure can be. Consider a set S whose only members are the people in Barack Obama's nuclear family: Barack, Michelle, Sasha, and Malia. If the argument I've just given is correct, then the structure of this set is exhaustively described by enumerating its membership relations. But you may think that the structure of the set is more complicated than this. After all, the Obamas stand in many relations to one another. Why shouldn't these relations be included in a description of the set's structure? Supposing we can include such relations, then some sets will have a rich enough structure to ground our knowledge of other mathematical objects.

The clearest consideration which prevents including these relations in a description of a sets structure is a familiar mereological one; namely, that a whole will, in general, stand in bear different properties and stand in different relations than the parts which make it up. So, for example, my right leg is a part of me. My right leg is (more or less) symmetrical to my left leg. From this we can't conclude that my whole body is symmetrical to my left leg though, because it isn't. Of course, some inferences of from part to whole are justified. For example, if I've burned my right leg, then we say that I have burned my body in doing so. But I don't think this is true for the case of sets and their members. At the very least, it isn't suggested by our practice of labeling mathematical objects. In particular, the fact that we distinguish sets from groups, ring, fields, and all the other algebraic structures which consist of a set plus some additional structure. If sets really had the relations which

their members do, then it would be superfluous to distinguish between, e.g., the set of integers and the ring of integers.

Obviously though, Maddy's notion of set is somewhat revisionary. Given that she takes impure sets to be spatiotemporally located, she might think that this explains how a set can inherit a richer structure from its members, since it has the same spatiotemporal properties as its members. If this is the case, then she might be able to ground our knowledge of more diverse branches of mathematics in this way. However, this route is only viable if she can defend her view that impure sets are spatiotemporal. We now consider whether this view is defensible.

2.2.1 Mathematical objects as spatial

In her discussion of whether numbers could be objects, Maddy writes the following:

> [B]ut if [the object] is to be singled out... as this object or that, it must have some additional properties... these additional properties will be superfluous to the object's numerical functioning... why should three have these superfluous properties and not some others? (1990, p.85)

The idea is that if a number is an object, it has to be some object. But if it is some object, then it will have non-numerical properties. But then it is possible that two objects could have the same numerical properties, but different non-numerical properties. Since they have different properties, they cannot both be identified with the number (given the transitivity of identity). But then which one is it? Arithmetic 'is mute' on the matter, to use Maddy's phrasing.

Just as arithmetic is mute on the matter of number's non-numerical properties, so too is set theory mute on their spatiotemporal properties. This is just a trivial consequence of the language in which we do set-theory: $\in$ is the only non-logical relation in the language. It is this sparseness which makes set-theory's success as a foundations for mathematics so remarkable, but it is also what prevents it from telling us anything about the spatiotemporal features of sets. It simply doesn't have the expressive power to do so.

If this is the case though, then Maddy's proposal that sets are in fact spatiotemporal has the jarring consequence that set-theory is an incomplete study of sets. To be more specific, there will be aspects of sets which cannot be studied by set-theory. Of course, Maddy would likely make the comparison to other sciences at this point. Human beings are, among other things, physical. They could therefore be accurately described as possessing certain physical properties. For example, one could give an aerodynamical account of how streamlined people are. This appears to describe an actually property of human beings. But anthropology, the study of people, has (I assume) nothing to say about their aerodynamical properties. Maddy could thus claim that, if the above argument about sets and spatiotemporal properties is valid, then we should also be able to claim that humans don't actually have aerodynamical properties. Since this is clearly wrong, we should instead conclude that a class of objects can have properties besides those described by the field which studies those objects.

While this response seems plausible, I don't think the analogy it is based on is quite right. Arguably, mathematical theories lay a claim to exhaustively describe their objects of study in a way that empirical sciences don't. To see this, compare what would happen if you asked an anthropologist about the aerodynamics of people with what would happen if you

asked a mathematician about the spatiotemporal properties of sets. The anthropologist, unless she were also a physicist, would likely respond by saying 'I don't know, you would have to ask a physicist about that'. The response from the mathematician would be much different. After some quizzical looks, she would probably tell you that you have simply misunderstood what a set is.

I would contend that cases like the former where a practitioner in the field would acknowledge that the question is well-formed but that they don't know the answer are naturalistically acceptable. On the other hand, if a practitioner in the field would tell you that you have misunderstood the nature of the thing they study, then your question is probably not naturalistically acceptable. If assuming the spatiotemporality of sets puts us in the second category, then we ought to deny that sets are spatiotemporal. Doing so, as we have already discussed, seems to preclude Maddy's ability to explain the diversity of mathematics. It would appear therefore that Maddy's version of a Weak Gödelian epistemology faces some insurmountable problems. In the next chapter, we present a structuralist version of Weak Gödelianism which is designed to solve these problems.

Chapter 3

A Structuralist Version of Weak Gödelian epistemology

You ask the name of this nature, it is spirit; you ask its office, it is that of an Angel. In as far as he exists, an Angel is a spirit; in as far as he acts, he is an Angel.

- St. Augustine of Hippo, *Serm. in Psal. 103*

3.1 Structuralist epistemology

As we saw in the previous chapter, Maddy's set-theoretic realist approach to Weak Gödelian epistemology faces serious problems. First, it is not apparent that the epistemology she introduces can account for the diversity of mathematical knowledge *as it is actually gained*. Additionally, positing that all mathematical objects besides sets are actually properties of

sets forces us to give a revisionary semantics for most of mathematics. Or at the very least, it undermines the epistemological advantage of the Weak Gödelian approach. After all, in ordinary cases perception justifies beliefs such as 'My wagon is red'. It is non-trivial to make the further claim that it also justifies beliefs like 'Redness is instantiated in this wagon'. But if the number 2 is actually just a property of sets, then the sentence '2 is prime' means something closer to 'Primeness is instantiated by any set which instantiates 2-ness'. Finally, her insistence that sets are spatiotemporal leads to issues.

Much of Maddy's motivation for adopting the number-property position comes from Benacerraf's *What Numbers could not be*(1965). The argument she considers is the familiar one, and runs as follows: Suppose that natural numbers are sets. Then they have to be *some* sets. Which sets then? Well, they could be, for example, $\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, ...$ These are the Zermelo ordinals. But the natural numbers could also be $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}...$ These are the von Neumann ordinals. However, the natural numbers cannot be both. If we are to declare that one of these assignments is the correct one, we should preferably be able to say why this is. Since both collections can play the role of the natural numbers equally well, arithmetic cannot decide between the two. Thus, we would need some extra arithmetical argument. But how could there be an extra-arithmetical argument which told us facts about the numbers? It seems as if there cannot be. We should therefore conclude that neither of these collections constitute the natural numbers. But we could run this argument with any arbitrary collection of sets with a rich enough structure to represent the natural numbers. So we can conclude generally that there is no correct identification of the natural numbers with sets.

This argument is then supposed to generalize to show that natural numbers cannot be

any object. To say that 'this object is the natural number n' we first need to be able to individuate that object. To do so, we will need to appeal to some non-numerical property. Since there are no arguments that can show what the *correct* non-numerical properties are for three, there can be no argument for any identification of the natural numbers with other objects whatsoever. So the natural numbers are not in fact objects at all.

Assuming that numbers are properties leads to unacceptable consequences, and yet the arguments just sketched seem to force us to take the number-property position. How are we to manage these conflicting pressures? Maddy hints at the following route, saying 'the second argument leaves room for the position that numbers are the sort of objects whose non-superfluous [i.e. numerical] properties are all the properties they have' (1990, p.86). She rejects this account because it 'flaunts ontological economy' and 'requires an account of the sort of 'abstraction' or 'association" which can account for how it only has numerical properties.

Hindsight gives us a perspective on this question which was unavailable to Maddy. At the time of her writing, she was correct in saying that no serious and sustained investigation of objects with no non-numerical properties had been made. However, given the development of the structuralist program in the late 1990s by Shapiro and Resnik, this is no longer the case. Seeing as there is now a well developed theory of such objects, the following project suggests itself: give a Weak Gödelian account which doesn't take mathematical objects to be properties of sets, but instead takes them to be structures. As we'll see, such a marriage of structuralism and Weak Gödelian epistemology is fruitful. The Weak Gödelian line does much to clarify the more obscure aspects of structuralist epistemology, while structuralism gives the Weak Gödelian an ontology which can avoid

the Benacerrafian hang-ups laid out above (the very task structuralism was created to do). Before we carry out this combination, we outline the structuralist position.

3.1.1 Structuralism and pattern recognition

Ante rem structuralism, also called Platonic structuralism, is the view that mathematics is the study of abstract, independent objects called structures, and that mathematical objects are places in those structures. Define a *system* to be a collection of objects with a number of relations which hold between those objects. A structure is the pattern which is found in different systems whose objects stand in the same relations to each other. To use a favorite example of Shapiro, the defending team on a baseball field consists of a collection of people, and these people stand in certain relations to each other (e.g. The second basemen is standing between the shortstop and the first basemen). Thus, we say the defending team is a system of objects. On the other hand, other baseball teams could arrange themselves in the same way on the field that this team has. The other team clearly constitutes a different system of objects, but we would say it instantiates the same structure. In this case, the structure both teams instantiate is a given baseball defense. Thus, structures are abstract objects which enter into instantiation relations with collections of particulars.

One may wonder how this is any different from Maddy's account. After all, number properties are also abstracta, and are instantiated by collections. At first glance, the only real difference is that Maddy's account specifies that the collections constituting systems must be sets, whereas the structuralist leaves it open what exactly the systems contain. But the main difference is that structures are, as their name suggests, more structured than number-properties. Specifically, structures contain positions. This allows us to say

more interesting things about in what way the system instantiates the structure. To continue with the baseball example, properties allow us to make general statements about the collection of players. For example, 'the team is playing a lefty-shift defense' would express such a general claim. But by with structures we can say something like, 'Corey Kluber is playing pitcher in that lefty-shift defense'. In this way we can not only discuss the general properties of the collection, but how individual members of the collection factor into those properties.

Turning to the mathematical, we can see how the *ante rem* view addresses the problem Benacceraf raises in *What Numbers could not be*. Consider again the Zermelo ordinals and the von Neumann ordinals. For the structuralist, the question of which of these two are *the* natural numbers is confused. The Zermelo ordinals can be a system which instantiates the natural number structure, but so can the von Neumann ordinals. But neither of them are the natural numbers- that identification is reserved for the natural numbers structure.

Similarly, it is incorrect to ask whether $\{\{\emptyset\}\}$ or $\{\emptyset, \{\emptyset\}\}$ is *the* number 2. Both of them can play the role of the number 2 in some system which instantiates the natural number structure, but neither is identical to the number 2. The number 2 is actually just the third position in the natural number structure. Thus, positions in mathematical structures can either be taken as an object in itself, or as a role played by other objects. As Shapiro writes in *Philosophy of Mathematics: Structure and Ontology* (hereafter PM), "In arithmetic, the natural numbers are objects, but in some other theories natural numbers are offices, occupied by other objects"(p. 11). This dual nature of positions in structures as both object and office allows us to sensibly interpret weaker mathematical theories in more powerful ones. Thus, though Augustine was talking about angels rather than the

identification of the natural numbers with the von Neumann ordinals, we can take his point nonetheless: you ask the name of this nature, it is set; you ask its office, it is that of a number.

So structuralism plausibly allows us to sidestep the ontological worries which trouble Maddy and Benacerraf. But how compelling is the corresponding epistemology? The structuralist[1] approach to epistemology is meant to explain how we can come to have knowledge of these abstract structures. By making mathematical objects into the kinds of things which enter into instantiation relations, the structuralist has taken a step in the right direction. However, we are owed more explanation than that. In order to provide this explanation, the standard structuralist line is to appeal to our capacity for pattern recognition. Shapiro writes:

> In part, our child starts to learn about cardinal structures [e.g. the 3-structure mentioned above] by ostensive definition. The parent points to a group of four objects, says "four", then points to a different group of four objects and repeats the exercise. Eventually, the child learns to recognize the pattern itself. (1997, p. 115)

This seems right, as far as it goes. But it's not clear how far that is. Specifically, it's not clear that it does any work in bolstering the structuralist picture. After all, the nominalist could object even at this early stage, and insist that the child has not learned to recognize "the pattern itself". All the child has learned to do, the nominalist might say, is to recognize physical systems of which it is correct to say 'this system exhibits the four-structure'. But this does not require us to reify the pattern itself.

[1]For the time being, by 'structuralist', I understand '*ante rem* structuralist'.

Ontological considerations

I think Shapiro appreciates the need for further ontological argument, because immediately after he finishes discussing the epistemological story, he turns his attention towards questions of object and identity. He takes his cue from Robert Kraut's 1980 article *Indiscernability and Ontology*, which itself elaborates and extends Geach's notion of relative identity. The discussion is extremely interesting. We start with the identity of indiscernibles. The notion of indiscernability is the standard one: two things are said to be indiscernible iff there does not exist a property s.t. the one thing possesses the property, but the other doesn't. We then make that the observation that for a language[2] which did not have any vocabulary to express the properties which differentiate to objects, a speaker of that language may identify those objects.

The point is illustrated by the following example: Consider an economist Milton who, while on the clock, speaks a restricted version of English called ML. ML is a language which does not have the necessary vocabulary to distinguish between two people who have the same income. Thus, for any predicate P in ML, if x and y have the same income, then P either applies to both, or P applies to neither. How would the world 'look' to our economist who speaks this language. Well, if he went for a stroll through the park for lunch and saw someone walk by who made $35,000 dollars annually, he might say 'Ah, look, it's $35,000!'. If he saw a different person with the same income walk by, he might say 'There's $35,000 again!'.

Of course, this story isn't quite believable. After all, it seems that, supposing our

[2]In what what follows I use the terms 'sublanguage' following Kraut and Shapiro, but one could just as easily imagine someone with less conceptual resources rather than less linguistic resources making the same judgements.

economist's senses all worked, he will innately possess the capacity for discerning people with different incomes. The story may become more believable if we suppose that our economist is actually a robot, Milton, who really possesses no more conceptual equipment than what is needed to determine income level.

Now consider a robot like our economics-bot, but instead of not being able to distinguish people with the same income, it is unable to distinguish collections with the same number of members. Shapiro writes:

> If our interpreter sees a clump of three trees, he might say, "For the number person, there is three." If three people walk by, he might remark, "There is three again."(PM, p. 121)

The general takeaway is this: The difference between objects and mere equivalence classes is dependent on the conceptual resources of the agent doing the distinguishing. What one language would call an object will be labeled as an equivalence class by a richer language which has the means to distinguish between members of that class. When restricting ourselves to an impoverished sublanguage, we could come to recognize patterns *as independent objects*. Thus, by utilizing this sublanguage procedure, we can gain knowledge about simple structures themselves.

One response might be to claim that such impoverished languages don't actually gain knowledge, seeing as they are simply incorrect to judge the equivalence class to be an object. But this would seem to rule out our ordinary knowledge of macroscopic objects. For example, imagine that your roommate has the perceptual and conceptual resources to distinguish between different collections of molecules which appear to us to be the

same object. When you look at your computer two days in a row and say 'This is my computer', your atom-collection roommate translates to his atom-collection distinguishing friends: 'What he really means is... this is a different collection of atoms in the arranged his computer-wise equivalence class'. You would be justified in saying that even though there were different collections of atoms, that doesn't mean there isn't just one computer there. Shapiro's point is that it is at least plausible that the number-structure person could levy the same complaint at you.

3.1.2 From simple structures to complex

Assuming that mathematics is 'the science of structures', it will have to account for structures that are more complex than the kind which we can come to know via direct perceptual access. Thus, a structuralist epistemology will need to explain how we can gain knowledge about these more complex structures.

Shapiro's account does so by appealing to two further capacities of ours: projection and description. The capacity for projection is essentially the capacity to recognize higher-order patterns. The example that Shapiro uses (PM, p. 118) is of a child learning about small cardinal structures by observing groups of strokes on a paper. The child is familiar with the I-pattern, the II-pattern, the III-pattern, and perhaps some others. Experimenting shows that by adding a stroke to the I-pattern yields the II-pattern, to the II-pattern yields the III-pattern etc. The child realizes that the different patterns can be arranged in order of how soon they come about from this procedure. Thus the child becomes familiar with the pattern I II III IIII. This pattern is clearly not a pattern of strokes, but a pattern of patterns of strokes. The child then realizes that, just as we can extend individual stroke patterns

by adding strokes, we can extend stroke-pattern patterns by adding stroke patterns. The child then projects this process, conceiving of patterns which they don't actually construct, but which they reason that could be constructed by continuing onward. This faculty for projection is how we come to have knowledge about larger finite cardinality structures, like the 10,001 structure exhibited by any system with 10,001 objects. Further, by projecting even further, the child may even be able to conceive of the ω pattern, exhibited by the natural number structure.

However, this only suffices to secure knowledge of finite (or at best, countably infinite) structures. To secure all other kinds of mathematical knowledge, we have to posit a third capacity: description. By giving coherent and categorical axioms, we thereby describe a structure. Understanding the description is to come to understand the structure. Shapiro gives the examples of someone describing a particular baseball defense by describing where each player stands, or of someone describing the structure of the US government by listing the various offices and the powers the office has, and how they relate to other offices. If we understand these descriptions, we gain knowledge of the structures they describe. In the mathematical case, we give a description by enumerating axioms. If we understand these axioms, we can come to have knowledge about the described structure, just as we can come to have knowledge of baseball defenses.

There are of course restraints on which descriptions yield knowledge of a structure. After all, we need some guarantee that the description will not be an empty one, and that it will be a unique one. Recognizing this, Shapiro writes:

> There are two requirements on an implicit definition. The first is that at least one structure satisfies the axioms. Call this the "existence condition". The

> second requirement is that at most one structure (up to isomorphism) is described. This is the "uniqueness condition"... Because isomorphism among systems, is sufficient for "same structure", a categorical theory characterizes a single structure if it characterizes anything at all. The tools for resolving the uniquenees requirement are thus found in mathematics itself. The same goes for the existence of structures... If Φ is a coherent sentence in a second-order language, then there is a structure that satisfies Φ. (PM, p.132-33)

So if a given description is categorical and coherent, then it actually describes a structure, at least on Shapiro's view. Thus, the question of determining which descriptions actually pick out a structure seems to be answered, at least provisionally (see the following subsection).

MacBride against structuralism

The jump from knowledge of simple structures to knowledge of large cardinal and infinite structures has come under attack, most notably in MacBride(2008). The criticism is two-pronged. The first prong attacks Shapiro's account of projection, while the second prong calls into question knowledge by description.

As far as projection is concerned, MacBride is happy to grant that we do come to initially come to understand the natural numbers in this way. But he notes that giving an account of the processes by which we form our beliefs is not the same as accounting for why those beliefs are warranted. So what the structuralist needs to do isn't merely to show that projection happens, but to argue for why it is justified. MacBride makes this point writing:

> It follows that if we are to come to an understanding of how the character of larger structures may be legitimately projected from the character of smaller ones, it is necessary to understand how the mathematical novice may pass from (1) particular knowledge that a given structure has a successor distinct from it, to (2) general knowledge that all structures, given or otherwise, have their own distinct successors.

MacBride then goes onto consider how our novice might go about justifying this transition from (1) to (2) by deriving (2) from (1), perhaps by appealing to some general axioms. This focus on *a priori* deductions of (2) from (1) is dialectically confusing for a couple of reasons. First and most trivially, MacBride hasn't formulated (1) and (2) in such a way that we should expect (2) to be derivable from (1). The jump from (1) to (2) isn't justified even if we assume the Peano axioms. (2) essentially states the successor axiom of PA, and so we shouldn't expect it to be derivable from other premises in PA.[3] This problem could most likely be easily gotten around by reformulating (1) and (2).

However, MacBride's focus on a derivational warrant for projection faces a more serious issue; namely, that it misses the mark of how Shapiro wanted to claim that (2) is justified. Shapiro writes in response:

> I never claimed that the transition to (2) is a matter of deduction from premises about pattern recognition, nor did I claim that (2) is apodeictic or even knowable *a priori*- and certainly not on the basis of the paltry results of pattern recognition. (2010, p. 142)

[3]Or, technically, (2) states the successor principle, which follows from the fact that the successor function is 1-1 and 0 has no successor.

Shapiro's preferred story is a holistic one. The child doing the projection adopts (2) as a hypothesis after confirming (1) a number of times. (2) proves to be a useful hypothesis, allowing us to learn all sorts of arithmetical facts. Arithmetic, in turn, proves to be an incredibly useful mathematical tool, and is used in a great deal of mathematics which informs us about the world. As (2) is seen to be necessary for more and more of our 'intellectual enterprise' and to yield good results, it is upgraded from hypothesis to known proposition, and eventually to a foundational axiom. There is no a priori deduction that justifies (2). Rather, (2) justifies itself by contributing to successful intellectual enterprise.

So much then for MacBride's objections to projection. What about description? Here, I think, MacBride's arguments get farther. To start off his attack on description, MacBride cites Frege's correspondence with Hilbert, in which Frege attacks Hilbert's view that establishing the consistency of a set of axioms establishes the existence of the structure which those axioms describe. Frege asks Hilbert what he thinks of these theological axioms: (i) All Gods are omnipotent (ii) All Gods are omnipresent (iii) there is at least one God. Should we accept that there is something which these axioms pick out, since they are consistent? But if we shouldn't, then what is the difference between consistent theological axioms and consistent mathematical axioms? Why does consistency entail existence in the latter case but not the former?

MacBride essentially poses the same question, asking "What guarantee do we have that even coherent categorical descriptions are not empty, failing to denote actually existing structures?"(2007, p. 162). In his response to this, Shapiro again takes a holistic line. There is no way to deduce the coherence principle from non-mathematical, self-evident premises. Rather, the coherence principle gains support from "being part of an overall

philosophy of mathematics which, I claim, well explains the enterprise"(2008, p. 147). So, again, MacBride appears to want some deductive justification of Shapiro's account, and Shapiro retorts that this demand misses the point.

But MacBride has another challenge. Granting that the coherence principle is true, MacBride demands an answer to "how do we know that descriptions are coherent and categorical?"(2007, p.162). Shapiro's response is the familiar anti-foundationalist line. We cannot base mathematical knowledge on any body of knowledge more secure, and so we are entitled to use mathematics to explain mathematical knowledge. This allows us to explicate coherence in terms of set-theoretic satisfiability (similar to how we explicate effective calculation in terms of Turing Computability). As Shapiro writes:

> In mathematics as practised, set theory (or something equivalent) is taken to be the ultimate court of appeal for existence questions. Doubts over whether a certain type of mathematical object exists are resolved by showing that objects of this type can be found or modeled in the set-theoretic hierarchy... the thesis that satisfiability is sufficient for existence underlies mathematical practice... Structuralists accept this presupposition and make use of it like everyone else. (PM, p.136)

In (2007), MacBride argues that, since knowledge of set-theory is also something that needs explaining, we cannot appeal to facts about set-theory to explain our knowledge of an axiom set's coherence. Doing so entails a vicious circularity. Shapiro denies this circularity in (2008), again emphasizing that he is not out to justify mathematics from non-mathematical premises. Because of this, he is entitled to help himself to set-theoretic

facts in spelling out his philosophical story.

This anti-foundationalist line does not work so well in this case. MacBride's challenge isn't to explain mathematical knowledge without appealing to mathematical facts. Rather, imitating Field, we assume that we do in fact have mathematical knowledge, and then try to explain how this happens according to our given philosophical story. It is no good for Shapiro to just accept that set-theoretic knowledge does happen, and then go on constructing a philosophical account given this. He is still required to show that his account doesn't make such knowledge impossible.

However, I take it that Shapiro has done so in all of the relevant cases, and would say that MacBride's second challenge goes too far. After all, what reason do we have to suspect that mathematicians know that the axioms of set-theory are coherent and categorical? Due to the Löwenheim-Skolem results, we actually know that the axioms of set theory aren't categorical (at least when stated using FOL). It is therefore unfair to ask Shapiro to show that his philosophical story can account for non-existent knowledge. So we conclude that the modes of projection and description have been adequately defended.

So the faculties of projection and description seem to hold up to some level of scrutiny. However, little has yet been said in defense of the first stratum of the structuralist epistemology: abstraction. From the point of view of the Weak Gödelian structuralist, this is unideal. After all, the capacity for abstraction is just as puzzling as projection or description. For example, consider the person who observes a system which instantiates the 3-structure. They then abstract from facts about the system to facts about the structure, and supposedly come to gain knowledge about that structure. But what fixes that viewing a 3- system and abstracting should yield knowledge about the 3-structure? If it amounts

to something like definite description, then there would have to be some mental act going from 'This 3 system is like this' to 'The structure which is exactly similar to this 3-system in the relevant ways is also like this'. But clearly there are ambiguities in such an act. What qualifies as exactly similar, and what are the relevant ways? These problems aren't insurmountable, but it would be simpler to not have to bridge this gap.

3.2 Cognitive basis of a Weak Gödelian argument

As we noted in Chapter 1, the Platonist should want to do more than counter any claims that a Weak Gödelian account is in principle impossible. They should aim to give a fairly explicit description of what such an account looks like. Ideally, such a description would follow Maddy in describing the cognitive processes responsible for ordinary perception, and then show how these processes could, at least theoretically, also form the basis for mathematical perception. In this section, I attempt such a description.

3.2.1 The psychophysical basis of ordinary perception

Recall the neurophysical account of perception given in Chapter 2: our perceptual system, like a number of cognitive systems, divides into neural layers. The neurons in a given layer detect are made to detect certain patterns of activity in the layer below, and their responses to these patterns are in turn detected by neurons in the layer above.[4] In the perceptual system, we spoke of three layers. First, a layer responsible for detecting sense stimulation.

[4]This talk of neural layers is obviously something of a simplifying assumption- in actuality, the connections involved are wildly complex, and don't even necessarily have the partial ordering structure which I've been presenting them as having. Nevertheless, the connections do, in general, display something like this layered connectivity, and computational models which make this assumption have proven very effective at duplicating the kinds of computations performed by the brain.

Second, a layer for detecting figures. Finally, a layer for detecting objects. As one might expect, the actual picture is more complicated than this. Since we are now adjudicating between competing Weak Gödelian accounts, we will need to go into a greater level of detail, so that we can see which account is better able to handle more specific empirical findings.

Let's begin with the sensory layer. For the sake of simplicity, let's also restrict ourselves to the case of visual perception. Visual perception begins with phototransduction. Phototransduction is the process by which light information is converted into chemical information by rod, cone, and photosensitive ganglion cells in the retina. This conversion is facilitated by the special chemical properties of these cells. The details of these properties are not essential, so we can summarize by describing their behavior in two different states: *stimulated* and *unstimulated*. The photoreceptor cells have a supply of the neurotransmitter *glutamate* which they use to communicate with the retinal nerves. In an unstimulated state, they release a constant amount of glutamate, which inhibits the retinal nerves.[5] However, a certain number of photons hitting the cell will cause it to enter a stimulated state. The stimulation of these photoreceptor cells by incoming photons causes a reaction which temporarily stops the release of glutamate. The absence of the inhibitory agent causes certain pathways to open in the walls of the cells of the retinal nerves, allowing positively charged ions from the intracellular fluid to flow in to the cells, causing a change in the cell's charge. This change triggers a chain reaction in the cells of the retinal nerves, with each cell changing charge, and in doing so releasing transmitters which cause subsequent cells to change charge as well. In this way, an electrochemical signal is sent along the retinal

[5]This is a unique case- in most cases, glutamate is an excitatory transmitter, not an inhibitory one.

nerves, to the optic nerve, and on to different parts of the brain. The optic nerve goes to more than one area of the brain, but the relevant area for visual perception is in thalamus (specifically, the lateral geniculate nucleus).[6] Here the brain carries out preprocessing of the visual information obtained. We can think of this processing in the following way: the information coming from the retina is encoded into bits, with a 1 representing 'light' and a 0 representing 'dark', with each bit associated with a given retinal cell.[7] We can again use the grid analogy from before when thinking about these retinal nerve cells: the retinal cells are those cells in the first layer which detect whether or not their section of the 'grid' is lit up.

Since the signals coming from the retina are not organized according to which part of the visual field they correspond to, the neurons in the thalamus perform this grouping. They divide the signals received into different packets, with different packets corresponding to different areas in each eye. In doing so, they facilitate the construction of a spatial representation at the next stage of processing. Metaphorically, we could think of the cortex as a campaign worker who has been tasked with making a map which shows public opinion of her candidate according to location. To help her make this map, she's been given a bunch of letters from people all over the country. The letters are all succinct. They either say, 'I like your candidate' or 'I dislike your candidate'. Her task would be greatly simplified if someone had already organized these letters by, say, postcode. That is essentially what the thalamus does. It takes the raw 'yes-no' data from the retina, and

[6] I say that this is the relevant destination because there is evidence that the other areas control eye reflexes and the sleep/wake cycle. These processes are not directly related to higher-level cognition.

[7] In actuality, the information encoded by any given retinal cell is slightly richer. In particular, there is evidence (give source) to suggest that retinal cells have already begun to emphasize edges between patches of light and dark. However, this simplification is fine for our purposes.

organizes it according to where it came from. From the thalamus, signals are then sent to the occipital lobe of the cerebral cortex for further processing. Here, the packets of information collected in the thalamus are processed individually by sets of neurons known as *hypercolumns*. These hypercolumns extract basic representational information from the packets, such as the relative location of edges, orientation, color, texture, etc.

All of the processes described above do the work which we have previously been attributing to one 'sensory' layer. It is not until we get to the deeper areas in the occipital lobe (specifically V2 and V4 of the visual cortex) that the work of the 'figure' layer begins. In areas V2 and V4 of the visual cortex, cell-assemblies detect certain arrangements of visual information represented by firing patterns of specific sets of hypercolumns. These arrangements will be largely invariant across different locations and sizes, meaning that a viewing of one scene from different viewpoints is likely to trigger the same response. The 'object' layer is found in the inferior temporal lobe/cortex (ITL), where cell-assemblies detect rotational invariance of those figures represented in V2 and V4 (i.e. figures represented in V2 and V4 which are the same apart from some degree of rotation will trigger the same response in the inferior temporal cortex).

This system of regions is collectively referred to as the ventral stream, and it is generally taken to play the principal role in object recognition.[8] This conclusion is supported by both fMRI data and the presentation of certain pathologies. For example, lesions in the inferior temporal lobe often cause specific visual agnosias, i.e. the inability to recognize objects visually (see Farah, 1990), while leaving other types of cognition unharmed. Additionally, in tasks which require switching attention from one object to another, fMRI data shows

[8] Or rather, the areas after the thalamus constitute the ventral stream, while the thalamus is said to feed into the ventral stream.

signals are fed back from the inferior temporal lobe to areas of the thalamus responsible for attention. In as far as the ability to focus our attention on a given object requires us to perceive the object, this correlation further implicates the inferior temporal lobe in object perception.

3.2.2 Philosophy from physiology

We can summarize the physical process of perception given above as follows. Photons hitting the retina cause an electrochemical signal to be sent to the thalamus. Here, the signals being received from the retina are organized according to location before being sent to the visual cortex in the occipital lobe. Then, basic representational information is extracted. This information is then further processed in V2 and V4 to provide less local, more general visual representations. Finally, these general representations are used by detectors in the ITL to identify spatially invariant objects, with the help of language processing centers in other areas of the temporal lobe.

This account goes a good way toward filling in the account which Maddy gives of regular object perception. But it also does a nice job of bringing out that feature of perception which underlies Maddy's first insight; namely, that perception of objects involves a significant amount of abstraction away from unnecessary details in our visual field. This point is also made by O'Reilly and Munakata:

> Instead of casting object recognition as a massive dynamic search problem, we can think of it in terms of a gradual sequence of transformations (operating in parallel) that emphasize certain distinctions and collapse across others. If the end result of this sequence of transformations retains sufficient distinctions to

> disambiguate different objects, but collapses across irrelevant differences produced by different viewing perspectives, then invariant object recognition has been achieved. (O'Reilly & Munakata 2000, p. 242)

So collapsing across differences in similar representations allows me to look at a chair, say 'there is a chair', walk to the other side of the room, and say 'there is that chair again' even though it does not look exactly the same to me. The general takeaway from our discussion of the physiology of perception should be this: in order for us to perceive objects at all, *we must ignore certain features of how those objects are presented to us.* When looking at my chair up close and farther away I don't posit two objects "chair near" and "chair far" with "chair near" being larger than "chair far". Rather, I posit one chair, and note that in the one case it looks larger than in the other. Thus, I ignore a feature of two different chair presentations in order to postulate one common object. The physiological account explains how this takes place.

The analogy with basic mathematical structures

With this story in place, we can return to the structuralist sublanguage procedure, the nominalist's objection, and, most happily, to our friends Milton and Carl. Let's again consider Milton object-judgements before and after applying a sublanguage procedure. Suppose Milton is standing around the office, and looks over by the water cooler, where he sees his colleague Joan standing. Using the full resources of his English vocabulary, Milton judges that 'Joan is standing by the water cooler'. Then he restricts himself to the sublanguage ML described above. Restricted to this language, he lacks the resources to differentiate Joan from any other person who is in the same income bracket as she is. Thus,

when he looks towards the water cooler this time, he judges '$50,000-$60,000 is standing by the watercooler.'

The physiological story told above makes the distinction between these two cases apparent. In both cases, visual information passes from the retina to the thalamus, then to V2 and V4 as before, where invariant representations of those objects are constructed. The difference arises during the processing in the ITL. Recall that the ITL is the area of the visual processing system which connects up to language processing areas in the temporal lobe. Recall further that this area plays a final, crucial step in object recognition, as evidenced by the fact that localized lesions in the ITL will hinder a subject's ability to perceive objects visually. These facts suggest the following account of Milton's behavior: In the normal case where Milton judges that Joan is standing by the water cooler, there will be activation in the visual stream, the ITL, and other language processing centers involved in evaluating sentences. In the sublanguage case, we will see similar activation in the visual stream and ITL, but would expect to see less activation in higher-language processing centers. We may further expect to see activation in the prefrontal cortex, which would be responsible for the executive functions controlling which kind of processing is taking place. By analogy, we would expect similar patterns of activation in our number person Carl, when he carries out the sublanguage procedure.

If the sublanguage account is a true one of how we can perceive simple structures, then these are the kinds of activation patterns we would expect to see in subjects who carry it out. Since I think the sublanguage account is the right account, I think our physiological story will be borne out. However, a final vindication for the sublanguage procedure could only come about with the help of a significant body of empirical evidence.

As far as I know, no fMRI studies have been carried out which provide exactly the right kind of evidence which the sublanguage account needs. In fact, there may be substantial theoretical and methodological issues which need to be resolved before cognitive neuroscience can fruitfully address the question of mathematical perception. Methodologically, studies are bound by the limits of the technology they use. In the case of fMRI, this places constraints on what mathematical tasks can be imaged. Tasks which take longer than a few seconds make it "harder to control for irrelevant intruding processes (e.g. mind wandering)"(Arsalidou et al., 2017). As a result, most of the studies in the literature only look at neural activation during simple tasks like judging which number of two is greater, or performing basic arithmetic operations. This prevents research on a number of philosophically interesting aspects of mathematical cognition, such as the acquisition of new concepts. Furthermore, ethical considerations prohibit extensive experimentation on younger children in the process of learning how to do math.

At a more theoretical level, there is a problem with identifying which processing functions are implemented by the different brain areas. In particular, due to the distributive nature of the processing which goes on in the brain, we cannot be certain that there is a 1-1 correspondence between activated areas and tasks performed. Because of this, it is often hard to tell whether a brain area is actually implementing the process we say it is. To clarify this point, consider the following example: suppose we ask a subject to perform a task T, and observe activations in areas A_1 and A_2 on the fMRI. We can conclude either 1)A_1 and A_2 process T jointly or 2) T actually consists of two separate tasks T_1 and T_2, each of which is processed by only one area. Such indeterminacies make interpreting fMRI data difficult. However, such problems may also be tractable to more refined experimen-

tation. It would therefore be too pessimistic to conclude that neuroimaging will never be able to provide compelling evidence in favor of or against the sublanguage account. We may not be there yet, but there is reason to hope that improvements in technology and experimental design could make it possible someday (perhaps soon).

Even if there is no direct evidence fMRI evidence for or against the sublanguage account, we may still be able to judge its plausibility from the results of related experiments. If these related studies find activation patterns similar to what we would expect if a subject were carrying out a sublanguage procedure, then this makes the sublanguage account more plausible.

The first finding we would expect, given a sublanguage account, would be a significant activation of visual processing centers when recognizing numerical patterns (whether counting or making comparative size judgements). There is a wealth of evidence to suggest that these areas are activated in such tasks. Studies by Piazza et al. (2007), Santens et al. (2010), and Roggeman et al. (2011) have all found significant activation in the visual stream. Further, this activation is similar to the activation patterns found in object recognition. This point is important enough to quote Roggeman et al. (2011) at length. They write:

> The processing gradient across occipitoparietal cortex [during numerosity recognition tasks] is similar... to gradients for processing objects more generally. These earlier studies observed gradients in the ventral visual stream that exhibited greater invariance for different stimulus properties (e.g., stimulus location) in a posterior-to-anterior direction. In the current study, a similar principle was described for the dorsal visual stream. In particular, to extract number, one

> must abstract across various irrelevant dimensions, such as object size and position... this study illustrates that a stimulus-dependent to stimulus-invariant gradient is a common organizational principle in the human brain, not only for object or word recognition in the ventral stream, but also for other stimulus properties in the dorsal processing stream.

This seems to be a crucial point in favor of the sublanguage account; *namely, that the same sequential discarding of irrelevant stimulus-information involved in object perception is also involved in recognizing number patterns.* The Weak Gödelian simply asks why we think of the one case as perception, but the other as some metaphysically occult process.

In addition to significant activation in the visual stream analogous to that seen during object perception, we would expect other areas of activation while implementing a sublanguage procedure. First, we would expect the activation of some center for mathematical concepts. This follows from the analogy to ordinary perception: we need to possess the concept of a chair before we can identify chairs. Initial studies seem to suggest that a region of the brain called the intraparietal sulcus (IPS) meets these requirements. As Dasterjadi et al. (2013) write:

> Our study provides the first intracranial electrophysiological data in humans supporting the important role of the IPS region in numerical cognition. Moreover, our observations suggested a surprisingly high degree of sparseness of activity in the numerically responsive IPS neuronal populations during the natural condition. Although these populations of neurons showed a high degree of selectivity for the experimental condition of numerosity, they exhibited a great

> sparseness of HFB response during the natural condition when the participants were engaged in activities such as attending to visual objects in the room, recalling past memories, grasping and reaching movements, or saccades?activities that are known to activate the LPC. The selective and sparse activity in the IPS neuronal populations provides direct evidence that these populations have a specific role in numerical cognition rather than being co-opted during various other parietal functions.

Together with the studies showing activation of the visual stream, the finding of a devoted semantic center for mathematical concepts strengthens the cognitive analogy with object perception. However, the sublanguage account doesn't just postulate the use of mathematical concepts. It also postulates that we restrict ourselves to only mathematical concepts. As far as activation patterns go, this should correlate with limited activation in other language processing centers of the brain.

The empirical results here are interesting, and generally support the sublanguage account. Almaric and Dehaene (2016) write:

> In our task, language areas were activated only transiently during the presentation of auditory statements, whether mathematical or nonmathematical. Rather, the activations that we observed during mathematical reflection occurred in areas previously associated with number coding in humans and other animals.

This would appear to support the idea that, when we do mathematics, we act only *qua* number person, restricting ourselves to our mathematical vocabulary. However, the

research only found that this result only consistently held of professional mathematicians. Laypeople were much more likely to utilize general linguistic areas of the brain when dealing with statements about higher mathematics (topology, algebra, etc.). Interestingly then, it would appear that trained mathematicians have the capacity to utilize something like a sublanguage procedure for higher math, whereas laypeople only utilize it for simple mathematical statements. Nevertheless, the overall findings support the sublanguage account.

Jones: Approximate Number System vs. Object Tracking System

In (2015) and (2018), Max Jones pursues an extension of Maddy's Weak Gödelian project just as I do in this thesis. But instead of moving from number-properties to structures, Jones follows Kitcher and makes the move to enumeration affordances. So according to Jones, we do really perceive numbers, and these numbers are just opportunities for us to count the objects which are numbered.

I'm not convinced that this move to affordances advances the dialectic in an important way. This is because the ecological theory of perception from which he draws the concept of 'perceiving affordances' is itself highly controversial. Most cognitive models require some sort of neural representation of what is perceived[9], but ecological theories of perception deny this. Instead, they claim that we perceive aspects of our environment directly, without having to represent them. To establish this though would require serious philosophical argument, especially in light of Fodor and Pylyshyn's criticism (1981).[10]To base a Weak Gödelian account on such a theory seems unhelpful then. No one doubts that objects are

[9]N.B. This isn't to say that what is perceived is the representation, so neuroscience is silent on the question of direct vs. indirect perception.

[10]For a detailed discussion of the shortcomings of the ecological theory, see Bruce *et al.* (2014)

given to us by perception, so to say that numbers are given to us in the same way does significant work in explaining how we come to know about them. But plenty of people doubt that affordances are given to us in perception, and so to say that numbers are given to us in the same way as affordances will do little to make mathematical epistemology less problematic. It just makes it problematic in a different way.

Despite this dialectical shortcoming, Jones' account is remarkable in its engagement with the cognitive psychology literature. One particularly important point which this engagement brings out is that there are in fact three cognitive systems which potentially play a role in our basic number judgements: the Approximate Number System (ANS), the Object Tracking System (OTS), and the Pattern Recognition System (PRS). The ANS is a system located in the parietal lobe which is responsible for making numerical judgements of an approximate nature. This approximate nature is brought out in size comparisons- for the ANS to be able to judge which of two numbers is greater, the two must differ by a given ratio. If they do not differ by this ratio, then the ANS will judge that they are the same size. In contrast, the OTS is a system located in the occipital and parietal lobes which is responsible for maintaining a representation of distinct objects within our visual field, with the capacity to do so for up to four objects at a time. The OTS underlies cardinality judgements for collections of up to 4 objects. Furthermore, comparisons processed by the OTS are more accurate than those in the ANS, even when scaling factors are taken into account. The PRS is something of a more distributed system, but it is also involved in subitising.

An interesting upshot from Jones' discussion is that it is not clear how great the role is that each of these systems play in numerical cognition. Jones' own conclusion is that the

role of the OTS and PRS in numerical cognition is greatly exaggerated. Given the account I've sketched, I should want to deny this. After all, there is significant overlap between the OTS and the ventral visual stream. Since I've claimed that the ventral visual stream underlies the perception of basic numerical structures, it would go against my view if the OTS didn't actually play a significant role in numerical judgements.

Jones has a few reasons for downplaying the role of the OTS. First, he claims that neither system explicitly represents number. In the case of the OTS, for example, all that is represented is the presence of a small number of discrete objects. But, as Jones says, 'a representation of threeness is not the same thing as three representations of oneness' (2015, p.61). Further, he expresses doubts that the research which implicates the OTS rather than ATS in numerical judgements for small collections 'comes from studies involving infants and young children that are still developing' (2015, p. 60). Lastly, given the activation of the IPS in numerical judgements, even for small collections, we should conclude that it is the ANS involved in these judgements, since 'neither the OTS nor the PRS are located in the hIPS' (ibid).

This reasoning can in fact be called into question. First, while it may be that the OTS by itself doesn't have the resources to represent number, the PRS combined with the OTS may be said to represent number structures. After all, there is evidence that the PRS can discern the presence of different numbers of objects in the OTS, regardless of their configuration (Madeler & Shebo, 1982). But if we have a system which detects the presence of n-objects, what is to prevent us from saying that we have representations of, for example, the basic 3-structure? While Jones has rejected structuralism as incompatible with a naturalist epistemology for mathematics (a rejection we'll address in the next chapter), we haven't

done so. Thus, we are free to acknowledge that the OTS and PRS do represent number in a meaningful way.

The concern that research supporting the role of the OTS and PRS in numerical cognition is too focused on immature subjects seems misguided. Yes, most of the research on the role of the OTS in numerical cognition *has to be* on immature subjects, because '[w]hen the counting principles are acquired... traces of the OTS in symbolic appear to disappear' (Piazza, 2010). So researchers are constrained to study children who have yet to master the counting principles. However, the concern seems misplaced in an argument in favor of the ANS, seeing as the ANS is also an immature system, which plays little role in mature numerical processing (ibid). Further, just because a system acts as our primary numerical processing system at a certain stage of development, doesn't mean that it is irrelevant to explaining mature mathematical knowledge.

Lastly, the idea that the OTS and PRS can be ruled out because they are not located in the hIPS is simply not consistent with the experiemental findings. Studies have shown (see Cipolotti, Butterworth & Denes (1991)) that patients with localized lesions in the hIPS are unable to process information about collections of greater than 4 objects, but were able to, with increased difficulty, make correct judgements of collections with 1-4 objects. This would seem to support that, contra Jones, some other system outside of the hIPS is involved in smaller numerical judgements given the possibility of completing such tasks with despite damage to the hIPS. However, given the increased difficulty in doing so, there would appear to be some interface between these systems and the hIPS.

So Jones arguments against the OTS and PRS as involved in numerical judgements seem to be inconclusive. Thus, the structuralist Weak Gödelian position would appear to be a live

empirical option. However, these kinds of considerations lead us to a general conclusion about the relationship between the empirical literature on numerical cognition and the epistemology of mathematics; namely, that the empirical literature is diverse and unsettled, so philosophical conclusions based on it must be cautious. In light of this, we should hedge the conclusions we made earlier. Instead of saying that the empirical literature supports the structuralist sublanguage account, we should say that the structuralist sublanguage account is consistent with the empirical literature *as it stands*. As the weight of opinion shifts or cements in the empirical literature, we should revise our epistemology accordingly. In the end, I suspect that the correct story will be much more complex than the sublanguage account as presented here. Nevertheless, the trajectory in the literature does not seem to indicate the empirical findings will not soon rule out an epistemology of structure analogous to an epistemology of object. We therefore maintain that the structuralist Weak Gödelian account fits into the broader scientific picture of numerical cognition, answering Linnebo's demand for an external explanation of mathematical knowledge.

3.2.3 Spatiality: An objection

The view as sketched faces a problem. We have argued that a 'number-person' could have perceptions of basic mathematical structures. Moreover, we have argued that actual people, when reasoning about number, carry out cognitive processes very similar to the sublanguage procedures typical of 'number-people'. In so far as they do so, they can be considered to perceive basic mathematical structures.

We now ask: are these structures spatial or non-spatial? Either answer causes problems. If the structures perceived are spatial, then it is not clear in what sense they are

mathematical. One of the main tenants of the structuralist framework we've been working in is that the only properties had by mathematical objects are their structural properties. Thus, the 2-place in the ordinal four-structure is completely characterized by being after the 1-place and before the 3-place and 4-place. It has no other properties, and stands in no other relations. But, if the structures perceived by the number-person (or ordinary person *qua* number person) are spatial, then they will have non-structural properties. This seems undesirable: the whole point of introducing a structuralist ontology was to yield a domain of mathematical objects to serve as the semantic values of mathematical terms. In as far as mathematics does nothing to describe the spatial properties of its objects, we would be forced to conclude that mathematics provides an incomplete description of the objects it studies. I take this to be an unacceptable result.

Further, if our perceptions of structures are spatial, then it seems like these perceptions will provide the wherewithal to posit multiple two-structures. Consider our number person looking at their desk, which has two books and two sheets of paper on it, as well as three pencils closer to the books. Foregoing all concepts but his numerical concepts, he looks at the books and says 'there is the cardinal two-structure'. He looks at the pens and say 'there is the cardinal two-structure again!'. Now he stands up, and looks at the desk from farther away. How many objects does he perceive? On the structuralist standpoint, there is only one cardinal two-structure, just as for Plato there is only one Form of the Beautiful. Because of this, he can only perceive one cardinal two-structure. But now we suppose that his perception of structures also gives him spatial information. In that case, he would be able to conclude both that 'The two-structure is close to the three-structure' and 'The two-structure is not close to the three-structure' depending on where he shifts

his attention to. But, since both of these things would be true at the same time, this seems like a contradiction. To avoid the contradiction, he posits that there are actually two two-structures there: one close and one not close. So he concludes there is more than one two-structure, contradicting the *ante rem* structuralist hypothesis. Taking this route would lead to a variant of *in rebus* structuralism, wherein there are separate, spatially located two-structures associated with all the different two membered collections.

Perhaps the move to make is to allow that mathematical structures are spatial, but that they are distributed in space. Thus, the two-structure is in both places, in the same way in which a lump of clay could be divided and moved to two different places. Of course, this would lead to the conclusion that the two-structure has spatial parts which do not correspond to positions in the structure. This is undesirable, as this would yield a fact about a mathematical object which would not be settled by the objects mathematical properties. The same argument could be made about temporal parts: If the position of the two-structure changes over time, then it has temporal parts which will not be described by mathematics. Again, this result cannot be accepted.

So there is significant impetus to deny that the objects of mathematical perception have spatial properties. However, to do so is also perilous, as it puts us in opposition to Kant. Famously, Kant held that space and times were forms of intuition. One important upshot of this is that if an object is given to us through intuition, then it must be spatio-temporal. Since Kant took perception to be a paradigmatic case of intuition, it seems likely that he would deny the possibility of perception of a non-spatial object, such as the two-structure.

Why should we hold that the objects of intuition, and thus perception, must be spatial? As Strawson argues, this follows from the fact that the objects of intuition are particulars,

and the further 'intimate connexion between the idea of the particular instance and those of spatial and temporal ordering: space and time are the forms of particularity' (1966, p.52). So the argument is as follows: intuited, therefore particular, therefore spatio-temporal.

But this understanding of the objects of intuition misjudges the relation between sensation and intuition. As I argued above, *what is particular at the level of object can just as easily be viewed as general at the level of sensation.* Intuition is only possible when sensations are grouped together as an object in the same way that '[e]xperience, empirical knowledge, is possible only when intuitions are brought under concepts, when empirical judgements are made' (ibid, p. 30). So the fact that the objects of intuition are particular cannot guarantee that they are spatial, because relative to sensation, they are not particular.[11] Once we introduce this relativity of particularity, we have logical space to resist Kant's claim that to be intuited implies spatiality. Additionally, if an object can be general at one level while still counting perceivable, what is to stop it from being general at both levels and being perceivable? If this is allowed though, then the Weak Gödelian account goes through. Thus, it would seem that nothing Kant says necessitates that to be perceived requires spatiality. Strawson also acknowledges this point, writing:

> The only other of the listed theses which is assumed rather than argued for by Kant is that which I have numbered (4), the spatiality thesis... From the truth that objects of our experience conceived of as existing independently of experience of them are in fact spatial objects, it does not seem to follow immediately that the spatial is the only mode in which we could conceive of

[11] I am not trying to hold up the intuited objects as a counterexample to Kant's stance that particular implies spatial. Obviously, a non-particular spatial doesn't count as a counterexample. Rather, I'm trying to show that his motivation for accepting this conditional isn't actually relevant to the conditional at all.

such objects. (ibid, p. 25)

What we have done is to assert that this possibility is actualized: there are objects of our experience which are not in fact spatial objects. Specifically, the basic mathematical structures which we perceive. Since, as Strawson points out, Kant has not actually given arguments against this, we are free to hold this view without fear of repudiation by Kant scholars. We can therefore answer the worry posed at the beginning of this subsection: Perceived mathematical structures are non-spatial.

So we have avoided the objection of spatiality. However, there is a deeper objection, again suggested in Kant's writings, which is not so easily dealt with. To begin, consider this summary from Strawson:

> [T]he doctrine that human intuition is essentially sensible intuition... amounts, as Kant explains, to the doctrine that our awareness, or perception, of an object requires that we be affected by the object. This in turn amounts to no more than that the existence of an object perceived is one existence and our perception of it another and distinct existence, which is dependent on the former, therefore, in a non-logical or causal sense, even though the correctness of the description of it as a perception of the object is logically dependent on the existence of the object.

This would imply the following argument against the Weak Gödelian account: To be perceived, an object must be sensed. To be sensed, an object must be causally efficacious. Mathematical objects are causally inert, therefore they cannot be sensed, therefore they cannot be perceived. Thus, a Weak Gödelian epistemology is a non-starter. This is perhaps

the most relevant and powerful argument against a Weak Gödelian epistemology, and we therefore address it at length in the next chapter.

Chapter 4

Information and Interaction

If the conclusions of the previous chapter are correct, then the neurophysical processes which give rise to ordinary object perception could, with little extension, give rise to perception of very basic mathematical structures. Thus, ordinary object perception and basic mathematical perception deserve comparable treatment. If we wish to avoid skepticism about ordinary cases of object perception, it therefore follows that we should accept basic mathematical object perception as being unproblematic.

HAVE YOU We have already dealt with methodological concerns about whether our argument is sufficiently normative. Specifically, we asked whether an analysis of the cognitive processes underlying mathematical cognition can explain why the beliefs we form in this way are warranted. As we saw, the warrant is second-hand. Assuming ordinary perceptual beliefs are warranted, we have insufficient reason to suppose that mathematical beliefs formed in an analogous way were unwarranted.

In response, the nominalist may try to restart the dialectic with the following objection:

The fact that the process of ordinary perception yields warranted beliefs but mathematical perception doesn't has nothing to do with a difference in process. Rather, it is due to a difference in object. Given that the two processes are so similar, we would expect them to warrant beliefs about similar objects. But since the objects of mathematical beliefs are so different from the objects of ordinary belief, we can deny that the latter kind of beliefs are warranted, even though they are formed by similar processes.

The motivation for this objection is seen in the Kantian argument sketched at the end of the previous chapter: We saw that we should deny the spatio-temporality of the mathematical structures which we perceive. But our senses are such that they are only affected by spatio-temporal objects. Since an object must affect our senses to be sensed, mathematical structures cannot be sensed. But, given the eminently plausible assumption that sensation is a necessary condition of perception, this rules out the perception of mathematical objects. Thus, a Weak Gödelian epistemology is also ruled out.

This response is simply another iteration of the Benacerraf problem. Given what mathematical objects are, how do we come to gain perceptual knowledge about them? And it seems like the problem has not lost any of its original force, despite this iteration.

If the Weak Gödelian is to make an answer to this reiteration, they will need to establish that the objects of mathematical perception can be perceived despite being neither spatial nor temporal. This is no mean feat. Lest things appear hopeless for the Weak Gödelian, here is the line I end up taking in response I defend: *the sensation associated with the perception of a given object needn't be a sensation of that same object.* In fact, in the case of mathematical perception, the sensation associated with perception of mathematical structures is the same as in the case of ordinary object percpetion.

4.1 Hart's Energy Argument and Maxwell's demon

Why think that the same neurophysical process can warrant belief about ordinary objects, but not mathematical structures? The intuitive answer is that ordinary objects can interact with the neurophysical mechanism, whereas mathematical structures cannot. As W.D. Hart puts the matter:

> Granted just conservation of energy, then, whatever your views on the mind-body problem, you must not deny that when you learn something about an object, there is a change in you. Granted conservation of energy, such a change can be accounted for only by some sort of transmission from, ultimately, your environment to, at least proximately, your brain. And I do not see how what you learned about that object can be *about* that object (rather than some other) unless at least part of the energy that changed your state came from that object.... Platonists like me may need a "physics" for abstract objects, a metaphysics if you will. (1977, p. 125)

On the other hand, we might object to the idea that (at least our present) science somehow rules out a naturalized epistemology of abstract objects. As Jaegwon Kim puts it in (1981), "the inference we must resist is 'Abstract; therefore, causally inert; therefore, unperceivable and unknowable"(p.347).Though Kim here is talking about abstract mathematical properties, I think we can resist the inference just as well in the case of mathematical objects. Just as Chapter 3 was intended to show that our best scientific understanding of object perception could also explain mathematical perception, our goal in this chapter is to show that our physics of ordinary objects need not be distinct from Hart's "physics of abstract

objects".

We begin with the argument which Hart sketches above. One crucial assumption made in this argument is that at least some of the energy which induces the change in the knower resulting in perceptual knowledge must come from the object perceived. For brevity, let's call this *Hart's Condition*. If this is the case, then perceptual knowledge of abstract objects, and in particular mathematical objects, is clearly ruled out. Worryingly for the Weak Gödelian, Hart's condition seems plausible. After all, in trying to construct cases where an object is perceived but no energy is transferred to the perceiver, we introduce a gap which could give rise to epistemic puzzles.

To see this, consider a perceiver P and an object O. In the normal case, in which Hart's condition is met, there is some channel by which energy is transferred from O to P, resulting in a change of state in P which we associate with knowledge acquisition. By assumption, the presentation of the energy in a certain fashion triggers some neurochemical process leading to the mental state corresponding to knowledge only when. But in a case where the energy resulting in P's change of state comes from some intermediate source besides O, we aren't guaranteed that the energy will be in a triggering state only if O is actually how P perceives it. To make these considerations more concrete, consider a woman looking at a tree by the lake. By looking at the tree directly, she comes to know that the tree is flowering. At least some of the energy responsible for her change in mental state must come from the tree (photons collide with the tree, this collision transfers energy and changes photons' direction, and these photons then hit the woman's retina, beginning the perceptual process). Assuming that she is not prone to hallucinate, the woman will only perceive the tree as flowering if the tree is actually flowering. In contrast, consider

the instance where the woman comes to know that the tree is flowering by looking at the water. In this case, we can imagine that the energy which causes the state change in the woman comes solely from the water (the photons collide with the water, transferring all of the energy derived from the tree to the water, while absorbing 'new' energy from the water).[1] However, it is easy to think of cases where these same conditions could be simulated, even if the tree weren't flowering. For example, a mad scientist could project images of flowers onto the surface of the water to augment the tree-reflection image. In this case, the projected light could give rise to the same 'water energy' signature as in the case of veridical perception, but in this case energy signature is not indicative of the tree's flowering. Such considerations lead us to endorse Hart's condition.

The Weak Gödelian has good grounds to reject this argument for Hart's Condition. It is not apparent that any of the various kinds of energy behave as the above argument describes. In particular, it isn't obvious that energy is indexed according to any object besides the object which possesses it. To be more precise, we note that energy is a measure quantity associated with either a system or an object, just as mass is a measure quantity for objects. Thus, the X Joules of energy possessed by an object can be said to be that object's energy. But it doesn't seem like we can say that Y of those X Joules are 'tree energy' or 'water energy', because the object gained Y Joules when it interacted with the tree or the water. This is because, were that possible, we would be liable to a regress. After all, 'tree energy' wouldn't ultimately be tree energy. The tree would have received that energy from some source (maybe the sun, groundwater etc.) and in turn, that source's

[1] I don't actually know whether this is a physical possibility. It seems plausible to me that the cases where the photon loses all of the 'tree energy' would correspond exactly to the cases in which the tree's reflection is not visible in the water. That would be a point in Hart's favor though anyway though, so we can grant it for the sake of charity.

energy will have come from somewhere, and so on. It would therefore be difficult to explain why the 'tree energy' made the perceptual knowledge about the tree rather than about some other object in the tree energy's hereditary history.

In short, talk of energy 'coming from' objects is not sufficiently intelligible. But let's suppose that talk of energy 'coming from' an object could be satisfactorily spelled out. Even in that case, what is the motivation for holding that the transfer of energy has anything to do with the 'aboutness' of our mental states. That is to say, why does the energy which brings about my mental state need to, at least partially, come from the intentional object of that same mental state? Suppose I am looking at a picture of Barack Obama given to me by a photographer I know, and on the basis of the photo come to the conclusion 'Barack Obama has white teeth'. On Hart's view, one of two things is happening in this case. First, it might be that my knowledge in this case is not 'about' Barack Obama, but about this picture of him. Or else, it is the case that the picture of Obama is transferring energy to me which was previously had by Obama. The first conclusion seems too harsh: we come to think things about other people, places, things etc. all the time because of representations of those things which we see. Even though I, and many others, have never seen Obama in person, it seems natural to say that many of us know things about him.

On the other hand, what can be said to defend the idea that all representations convey energy from their original source? It would seem that this is not something which can or should be settled *a priori*. But Hart has given us no empirical considerations in favor of this view. Lacking such considerations, it seems open to the Weak Gödelian to reject Hart's Condition and defend the possibility of perceptual knowledge of mathematical objects.

4.1.1 Refining Hart's Condition

At this point, Hart's condition has not been defended well enough to pose a serious threat to the Weak Gödelian view. However, the motivation underlying Hart's condition is a powerful one: coming to know about an object involves an interaction between us and that object. If we want to make this interaction naturalistically scrutable, there should be some sort of physical medium which underlies it. Though we have seen that proposing energy as this physical medium leads to problems, perhaps there is a way to reformulate Hart's condition which respects the original motivation, but doesn't run into the same problems of unintelligibility.

One promising line to pursue is to consider the relation between energy and information. To illustrate this connection, we can start by looking at a thought experiment due to James Clerk Maxwell (Maxwell, 1872). The thought experiment runs as follows: take a container, and divide it into two sections, A and B. Cut a small trap door into this divider, and leave it open. Fill the container with a set amount of an ideal gas. Leaving the trap door open for a time, allow the gas within to evolve into a state of thermodynamic equilibrium (that it will do so is guaranteed by the second law). Now close the door in the divider.

At this stage, the two sections of the container are the same temperature. Maxwell then has us consider a demon operating the trap door in the divider. The demon has superhuman (though perhaps not supernatural) measurement abilities. By using these abilities, it is able to open the trap door as faster, higher energy particles from A fly towards it, and then close it again when slower, low energy particles from A threaten to go through. In doing so, the demon raises the temperature of B and lowers the temperature

of A, without expenditure of work. This violates the second law, which requires that a system will not move away from equilibrium unless some work is done on the system.

One response to this antinomy is to claim that this process is problematic only if we consider the demon to be separate from the physical system. However, once we include the demon, or any intelligent agent capable of performing measurements, in one of these systems, we can no longer conclude that the system violates the second law. The demon must make some measurement in order to know when to open or close the trap door. Thus, the demon is storing information. The resolution of the antinomy is to recognize that storing information is an essentially thermodynamic process. To see this, consider another container split into two sections, with no trap doors in the divider. Assume that in one of these sections is a single gas particle. This container can be said to hold 1 bit of information, as it could be used to communicate the answer to a yes/no question (where 'yes' is associated with the particle being in the first section, and 'no' being associated with the particles being in the second section). To 'erase' this bit of information, we could proceed as follows: replace one of the outer walls of the container with a piston, remove the partition, and then move the piston into the spot where the partition had been. The end result is the same regardless of which side the particle of gas was in, and thus the information is lost. However, the compression of the piston will require energy to overcome the pressure caused by the movement in the space of the gas particle. Thus, even in an ideal case, the erasure of information necessarily requires some energy. But, if the demon is assumed to be a physical creature, it will only be able to store a finite amount of information. If it is to continue functioning indefinitely then, it will have to erase some of the information it had previously stored. But this action will require the expenditure of

energy. This expenditure balances out the energy we had previously gained 'for free' from the operation of the trap door, and so the second law is satisfied.[2]

So there is a direct physical connection between energy and information. Given the additional relationship between information and knowledge, the nominalist might be inclined to try to reformulate Hart's condition in information-theoretic terms to see if the argument against perceptual mathematical knowledge can be resuscitated. As we'll see, I don't think this move to an information-theoretic Hart's Condition gets the nominalist any further, but I do think it gives us a more rigorous background to formulate Hart's argument in, and therefore makes the nature of the dialectic more perspicuous. With this end in mind, we now turn to a presentation of some of the basic definitions and results of information theory.

4.2 Physical information

Consider the rolling of a fair die.[3] This constitutes an event E, the outcome of which we are uncertain about. However, we do know the likelihood of the different outcomes ahead of time. Because of this, we can describe our expectations by means of a probability space X. X consists of a triple $(\Omega, \mathcal{B}, \mathcal{P})$, where Ω is the set of possible outcomes of E, $\mathcal{B}$ is a

[2]This is a simplified presentation of some conclusions which have emerged from a rich literature on the relationship between information processing and thermodynamics. Inevitably, some of the subtleties in this literature have been neglected. In particular, the debate between which operations on information have energy requirements. Szilard (1929), Gabor(19xx), and Brillouin(195x) have argued that information acquisition has intrinsic energy cost. On the other hand, Landauer (1961), Penrose (19xx), and Bennett (19xx) hold that information erasure is the process with the intrinsic cost. These distinctions are inessential to our discussion of mathematical epistemology— what is important for our debate is that information processing is related to energy, not how it is related. For a more detailed overview of the information thermodynamics literature, see Maroney (2009).

[3]In what follows, much of my treatment has been drawn from Jones and Jones(2000) and Cover and Thomas(2006).

σ-algebra on Ω representing an event space, and $\mathcal{P}$ is a function mapping $\mathcal{B} \to [0,1]$ such that (i) For $b \in \mathcal{B}, P(b) \geq 0$ (ii) $\sum_{x \in \Omega} = 1$ (iii) If b_1, b_2,...$b_n \in \mathcal{B}$ are pairwise disjoint, then $P(\bigcup_{i=1}^{i=n} b_i) = \sum_{i=1}^{i=n} P(b_i)$.

It may be useful for various reasons to be able to articulate how much uncertainty is associated with the event E ahead of time. Therefore, it would be nice to be able to define a quantity in our model which acts as the formal analogue of uncertainty. Such a quantity should satisfy three criteria:

1. The uncertainty should be at its highest when all the possible outcomes are equally likely.

2. The uncertainty associated with event A and B should be equal to the uncertainty associated with event A plus the uncertainty associated with event B.

3. The addition of impossible outcomes to the probability space should not increase the uncertainty

It can be shown that these three conditions yield a unique definition[4] for the uncertainty H associated with a probability space X given as follows:

$$H(X) = -\sum_{x \in \Omega} p(x) \log_2 p(x)$$

The value H is also sometimes referred to as the entropy of the probability space, given its formal similarity to Gibbs's formula for the entropy of a physical system.[5] By considering a couple of toy examples, we can get a better sense of how this value satisfies

[4]Or rather, unique up to a multiplicative constant.
[5]Compare to Gibbs Entropy: $S = -k_b \sum p(x) \log p(x)$

conditions 1-3 set out above, and thus is a good representation of the amount of uncertainty. So let's take the case of a fair die. In this case, $\Omega = \{1, 2, 3, 4, 5, 6\}$. For each $x \in \Omega, P(x) = \frac{1}{6}$. Plugging into our formula for $H(X)$, we get:

$$\begin{aligned} H(X) &= -\sum_{x \in \chi} p(x) \log_2 p(x) \\ &= -\frac{1}{6}\log_2(\frac{1}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) - \frac{1}{6}\log_2(\frac{1}{6}) \\ &= -6(\frac{1}{6}\log_2(\frac{1}{6})) \\ &= \log_2(6) \\ &\approx 2.58 \text{ bits} \end{aligned}$$

The quantity is given in bits because we were using log base 2 in our calculations. Thus, we see that an average roll of a fair die yields 2.58 bits of information.

Contrast this process with the throwing of a die which lands on 3 half the time, and 1,2,4,5, and 6 are all equally likely. Supposing we know the die is unfair, we are more certain about the outcome. Thus, we would expect the entropy of the process to be lower. We confirm this by calculating:

$$
\begin{aligned}
H(X) &= -\sum_{x \in \chi} p(x) \log_2 p(x) \\
&= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{10}\log_2(\frac{1}{10}) - \frac{1}{10}\log_2(\frac{1}{10}) - \frac{1}{10}\log_2(\frac{1}{10}) - \frac{1}{10}\log_2(\frac{1}{10}) - \frac{1}{10}\log_2(\frac{1}{10}) \\
&= -\frac{1}{2}\log_2(\frac{1}{2}) - \frac{1}{2}\log_2(\frac{1}{10}) \\
&\approx 2.16 \text{ bits}
\end{aligned}
$$

So when the discrepancy between the probabilities of the outcomes is higher, the uncertainty associated with the event is lower, which accords with condition 1. What about condition 2? For this, consider, instead of rolling a fair die, the flipping of two fair coins. Suppose that the flips are independent. Our set of possible outcomes now looks like the following $\{\{HH\}, \{TT\}, \{HT\}, \{TH\}\}$, with each having a probability of $\frac{1}{4}$. It is easy to see that the uncertainty associated with this space is 2 bits, since $\log_2 4 = 2$. Since it is also easy to see that the uncertainty associated with the flip of a single coin is 1 bit, we have evidence that uncertainty is additive in the way required by condition 2.

Condition 3 is fairly trivial, but note that its demonstration requires a limit argument, given that H(X) is undefined if X contains any truly impossible outcomes. If X' is a probability space with 1 additional possible outcome v, then it would be straightforward to show that H(X') approaches H(X) as $p(v)$ approaches 0.

4.2.1 Mutual information: Hart's Condition Revisited

So much for quantifying uncertainty. Having done so, quantifying information is very easy. The amount of information yielded by a given outcome is equal to the uncertainty previously associated with that outcome. Thus, the quantity H(X) also measures the average information yielded by a process, and the information yielded by a particular outcome x is given by -$\log_2 p(x)$.

One intuitive idea we have about information is that it can be transferred, just as heat can be transferred between two systems. Cover and Thomas put the matter as follows:

> What do we mean when we say that A communicates with B? We mean that the physical acts of A have induced a desired physical state in B. This transfer of information is a physical process (2006, p. 183)

This communication takes place via an information channel. An information channel is a system which includes a sender, a receiver, and a transmission medium. The sender encodes a character of a certain alphabet for transmission, sends it through the medium, and then the receiver decodes that character. We can capture this notion of communication using the formalism of entropy as follows: take two random variables X and Y. Consider their joint probability distribution for these variables $p(x, y)$. If X and Y are independent variables, then $p(x, y)$ is just $p(x)p(y)$. Furthermore, if they are independent, then we would expect X to contain no information about Y. For example, if the flashing of a light in one room were independent of the flashing of a light in the other room, intuitively the flashing of the one contains no information about the other. We therefore want the mutual information of X and Y to be zero. We give the mutual information between two events

as follows:[6]

$$I(X:Y) = \sum_{x \in \chi} \sum_{y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

On the basic level, this matches well our intuitive thoughts. In the case where $p(x,y) = p(x)p(y)$, $I(X:Y)$ is 0, which is as it should be. In the case where we know the probability of y given x (such as in the case where we know the reliability of a given channel), then information about $p(x)$ can be used to determine $p(x,y)$. Thus, if we know about the properties of a transmission channel, and the probability distribution of a random variable on one end, we can compute the mutual information with the variable on the other end. So we are also able to handle the concept of transferring information within this formalism.

Finally then, we have a robust enough concept to revitalize Hart's argument in section 4.1. To begin, we note that sensation is a necessary condition for perception. But our sensory systems are the kind of system which lends itself to information-theoretic analysis.[7] For example, in the case of vision, we can ask about the joint probability of a nerve cell in the retina firing and the presence of an illuminated object in a subject's vicinity. Assuming the subject's vision is reliable, i.e. there retinal nerve cells have a low probability of firing when there is no object present and high probability of firing when there is an object present, then the information conveyed by the firing of a retinal nerve cell will be about 1 bit.[8] And though a complete information-theoretic description of vision would be far more complicated, the principle would be the same.

[6]For a discussion of how this formula relates to and is derived from Shannon entropy, see Jones and Jones (2000, Section 4.6).

[7]In fact, such analysis has led to interesting insights into the development of our sensory systems. For example, (Delgado-Bonal and Martin-Torres, 2016) shows that the human vision system adapted to be sensitive to those wavelengths of radiation which maximize information gain given earth's atmospheric composition.

[8]In the ideal case where the nerve cell fires iff an object is present, the mutual information is exactly 1.

But what about the case of mathematical perception? Again, we suppose that sensation is a necessary condition for perception. Focusing on vision, we can ask what the mutual information is between, say, 2's being prime and the firing of certain nerve cells in the retina. Given that the two are independent, the joint probability will be equal to the product of the marginal probabilities.[9] Thus, the mutual information between the two is equal to zero. Since '[t]he mutual information... represents how much of the information in the input $\mathcal{A}$ is emerging in the output $\mathcal{B}$',[10]we conclude that vision yields no information about the primeness of 2. Given that our examples are arbitrary— all mathematical truths are true necessarily, we could have selected any and run the same argument— we conclude generally that our senses can garner no information about the mathematical. On the supposition that sensation is necessary for perception, we conclude that there can be no perceptual knowledge of mathematics.

4.3 A Weak Gödelian reply

So we have at least done some work on Hart's behalf of making his argument more intelligible. Rather than talk about the energy uptake, we could now phrase Hart's condition as follows: There can be no perceptual knowledge about a thing unless there is an information channel from that thing to the knower's sensory system. Given that mutual information is required for the presence of a non-trivial information channel, the lack of mutual information between mathematical objects and sensory responses rules out the perception of mathematical objects. This obviates the comparison between ordinary object percep-

[9] i.e. $p(x, y) = p(x) * p(y)$

[10] Jones and Jones, 2000

tion and numerical cognition made in Chapter 3, and derails the broader Weak Gödelian project.

There is much to recommend this way of approaching the problem. First, it is a rather economical mode of presentation. After all, framing the issue in terms of joint probability captures both the necessity of mathematical facts, and the abstract nature of the mathematical objects which underlie them. In short, it unifies the two biggest puzzles for mathematical epistemology into one information-theoretic formalism. It also makes explicit an analogy which is sometimes seen in the literature. Recall Hart's talk of 'some sort of transmission' quoted above, or MacBride's mention of 'signals' coming from the mathematical realm. The parallels with information theory were already implicitly present, so it seems worthwhile to bring them to light and examine how extensive they really are.

The Weak Gödelian must acknowledge the force of this argument, but I don't think they need to admit defeat yet. Because while information has proven to be a useful concept for the theory of communication, it isn't clear what the relationship is between information and knowledge. Clearly, there is some connection. After all, we say that someone is well-informed about a topic if she knows a lot about it. Further, we can come to know new facts by gaining new information. Although these considerations suggest that possessing information plays some role in the acquisition of knowledge, there is still a fair bit of ambiguity in this role, and in the concept of information generally.

One crucial ambiguity in the notion of information is brought out in the following scenario. Consider two friends, Billy and Bobbi, who are at a football game. Billy gets up and goes to buy some snacks. While he is at the snack bar, the results of a different

game in the league are announced. Billy hears the announcement through the speaker at the food stand, while Bobbi hears it from the speaker by her seat. This announcement informs both Billy and Bobbi of the result. We now ask: have Billy and Bobbi gained the same information, or different information? On the one hand, we might think they have gained the same information; namely the information *that* Team C beat Team W. On the other hand, the channels through which the information travelled seems to be different in both cases. At the very least, one channel would contain a wire going to the speaker at the food stand and the other would contain the other wire going to the speaker where Bobbi is sitting. Assuming that different information channels carry different information, there is reason to think that Billy and Bobbi have gained different information.

Considerations like these lead us to draw a distinction between two different kinds of information: semantic information and physical information. Physical information is the kind described by the entropy formalism given above, and which operates at the level of signals. We therefore say that Billy and Bobbi gained different physical information when they heard the announcement, because the announcement was conveyed to them through different signal media. Semantic information is the kind of information we gain when we receive a message that is meaningful to us (thus the moniker *semantic*); it is information 'about' something, information which has an intention. In our example, Billy and Bobbi gained the same semantic information about the result of the game, despite gaining separate physical information. One useful way to think about the difference is to note that physical information is quantitative, whereas semantic information is qualitative. As physical information is essentially a measure of diversity among a given array of signals, it describes how much can be said using that array. Semantic information is concerned

with what is actually said when that array is interpreted.

The relationship between physical information and semantic information has been discussed at least since information theory's inception, with Shannon flagging the distinction in his seminal 1949 paper "The Mathematical Theory of Communication". Philosophers have been treating the topic for almost as long: Bar-Hillel and Carnap (1953), Dretske (1981), and more recently Floridi (2011) have all discussed what the overlap is between the qualitative and quantitative notions of information.

However, the conceptual relation between semantic information and physical information explored in this literature is slightly orthogonal to our concerns. The most relevant discussion is found in Chapter 4 of Floridi (2011), where he discusses something like the following hypothesis:[11]

Strong Ontological Condition: There can be no semantic information without physical information.

Note that this condition answers the question of *whether* physical information must be present for there to be semantic information. The Weak Gödelian could proceed by denying that we need to gain physical information to gain semantic information about mathematical facts. But her position would be stronger if she could grant that physical information must be present for there to be semantic information. I believe such an option is open to her. Rather than asking whether physical information is necessary for semantic information, the Weak Gödelian should concern herself with *which* physical information specifically

[11] I have slightly altered the wording of Floridi's condition, since he states it in terms of his own metatheoretic account of information. It wouldn't be worth introducing this new terminology, since we are moving on to different issues directly afterwards, but see p.90 of Floridi (2011) for interesting discussion.

can be associated with some specific semantic information. If she can show that there is some relatively unproblematic physical information which can plausibly be taken to yield semantic information about mathematics, then there is no need to make a Weak Gödelian view conditional upon views about the nature of information. Such unconditionality is preferable.

Nominalist replies

It might seem that which physical information can yield some semantic information can be read off from the entropy formalism. The entropy is measure on a probability space, and this probability space in turn describes a physical process. We might say then that any semantic information associated with this physical information can only concern those things which feature in the physical process described.[12]

To respond to this view, we recall Maddy's second insight that we only partially sense the objects we perceive. So in an information-theoretic account of sensation, the process associated with the physical information will only feature parts of the object perceived. In the case of vision, this is the slice of the object which is visible to the subject. If we stipulate that we can only gain semantic information about objects in this process, we will only be able to gain semantic information about parts of the object, not the object itself. So this requirement precludes us gaining semantic information about objects through perception. Since we obviously do gain semantic information about objects via perception, we should reject this requirement.

Another plausible connection that the nominalist might consider is a similarity between

[12]I'm intentionally leaving open what objects exactly 'feature in' a physical process.

semantic and physical information with respect to probability. To see what is meant by a 'similarity' in probability, consider the following example: My friend and I are on two ends of a telegram wire, and she is flipping a coin and sending me the results. Thus, there is both physical information transfer going on, and semantic information transfer going on. Suppose that she uses one dot to communicate that her flip has landed on heads, and one dash to communicate that her flip has landed on tails. In this scenario, there are actually two probability distributions. One has the set of possible outcomes {Heads, Tails} and the usual probability function $p_1(x)$. The other has the set of possible outcomes {Dot, Dash} and the probability function $p_2(x)$. If my friend reliably sends dot when she gets heads, and dash when she gets tails, then $p_1(Heads) = .5 \approx p_2(Dot)$ and $p_1(Tails) = .5 \approx p_2(Dash)$.[13] Thus, the probability that the truth conditions for the sentence 'My friend has flipped a coin which landed on heads' obtain is very close to the probability that I will receive a dot along the wire.

The nominalist could give a general version of this stipulation as follows: Let PI be some physical information, which is correlated with the entropy of a physical process PP. Let SI be the semantic information conveyed by a sentence S. We stipulate that for PI to embody SI, the mutual information between PP and the truth conditions for S has to be non-zero. That is to say, the joint probability distribution for the outcomes of PP and the truth conditions for S has to represent some dependence between the two. To put it simply, there has to be some outcome of PP which is likely to happen just in case the truth conditions for S obtain, and some outcomes which are likely to obtain just in case they don't.

[13] We use approximation here to leave open the chance that my friend sometimes sends the wrong signal, but usually sends the right one.

If this stipulation is accepted though, then we seem to be back to ruling out mathematical perception. After all, there can be no physical process which both transmits physical information *and* is coordinated with the truth conditions of a certain mathematical fact in this way. To have a non-zero information content, the physical process must contain some uncertainty. But the closer the coordination between the physical process and the truth conditions of the mathematical statement, the less uncertainty there is in the physical process, with the limit case being when the physical process has the same outcome every time— that outcome associated with the truth condition of the statement (supposing it to be a true statement).

But the Weak Gödelian can push back. First, it would seem that the nominalist doesn't do well to acknowledge that there are truth conditions for mathematical truths which obtain necessarily. Doing so at least appears to implicitly depend on the idea that there is a realm of necessary mathematical objects which make our mathematical statements true. But this would just to concede the Platonist's point. However, upon further inspection, this is just the reliability argument specialized to attack the Weak Gödelian picture. The nominalist grants that there are these mind independent abstract objects, but then seeks to show that there must be some transfer of information which would underlie the sensation. Barring this possibility, there is no plausibility to the Weak Gödelian point.

The real issue, I think, lies in the following inference: a mathematical truth is necessary, therefore its truth conditions have a probability of 1. This inference only makes sense if we are using some objective interpretation of probability. Briefly, an objective interpretation of probability says that probability distributions represent the way things actually are in the world. For example, on the propensity interpretation of probability, processes have a

certain propensity to result in certain outcomes. Thus, a dice has the propensity to land with the '1' side up exactly one-sixth of the time. In the case of the truth conditions for a mathematical fact obtaining, these will obtain in every instance, and so should be assigned a probability of 1 on this view.

But, going back to an argument by E.T. Jaynes (1957), there is an established tradition of viewing the probabilities involved in information theory as subjective probabilities. Such probabilities are not interpreted as telling us about any sort of objective frequencies or propensities, but rather as describing our evidence or credence that a certain event will take place. On this interpretation though, the motivation to assign a probability of 1 to the truth conditions of an undecided mathematical sentence is missing. If the probability of the truth condition isn't necessarily 1, the nominalist's stipulation doesn't hold any water.

Interaction and Indispensability

So by making the move to subjective probability distributions, the Weak Gödelian can defuse the similarity stipulation. The nominalist could object to this employment of subjective information as follows: Consider our pre-theoretic notion of information. Is there any aspect of this notion which is captured adequately by a theoretical framework in which one thing can convey information about another thing without having somehow interacted with the first thing? If there is not, and it appears that the move to subjective probability yields such a framework, the move to subjective probability should be rejected.

Since their defense seems to rely on subjective probabilities, the Weak Gödelian must resist this move. I think there are actually principled reasons to do so. The most compelling reason has to do with what qualifies as an 'interaction'. Echoing an ongoing theme,

precisifying interaction yields a dilemma: either we define interaction in a way which rules out us gaining semantic information about ordinary objects, or we define it in a way which admits of gaining semantic information about mathematical objects as well.

For the first horn, we take a narrow view of interaction, and claim that two objects can be said to interact only if some physical process or chain of processes connects them. This precisification, along with the interaction requirement, entails that the only objects which one can gain semantic information about are those which the carriers of the physical information are connected to by a chain of physical processes. In other words, the nuggets carrying the physical information need to have a traceable link of physical connections to an object O for that physical information to correctly convey semantic information about O.

However, this view is too strict, and would rule out gaining semantic information about ordinary objects. To see this, consider a chair. The chair is constituted by an aggregate of different particles. We again wield Maddy's second insight: we never directly observe all the particles which make up this aggregate. Nevertheless, we take ourselves to perceive the chair constituted by the whole aggregate. Does this aggregate physically relate to the particles which make it up? It is not clear that it does. If it did, then it would stand to reason that the history of the universe would have been different, even in some minor way, if there had only been particles, and not aggregates of particles. That seems wrong though: what could the presence of the aggregate add? But then if the aggregate doesn't have this physical relation to the particles which make it up, then it seems like the chain of interaction is lost. Thus, we can have no semantic information about this chair. Since we can derive semantic information about chairs from perception, it would seem that this

notion of interaction is inadequate. Thus, we should conclude that the idea of physical interaction is not contained in our pre-theoretic notion of information.

Suppose, in order to allow for semantic information about chairs, we broaden our idea of interaction to include cases where two objects stand in a metaphysical relation to one another, in addition to physical ones. So the chair is constituted by the aggregate of particles (metaphysical relation) and the aggregate includes all the particles which make it up (metaphysical relation-assuming inclusion is so?), and the particles interact with the photons, etc. which convey the physical information to the subject. Thus, we have a chain of interactions which permits the subject to derive semantic information about the chair from the physical information.

But if we are allowing metaphysical relations to constitute interaction, then there is room for mathematical objects on this picture. To see this, consider the above example of

$$Chair \rightarrow aggregate \rightarrow particles \rightarrow photons$$

where each arrow represents a relation, either physical or metaphysical. Now suppose that instead of one chair, there are three. We can construct the following chain just as easily:

$$3\ structure \rightarrow chairs \rightarrow aggregates \rightarrow particles \rightarrow photons$$

In this case, the first arrow represents instantiation, and the second arrow represents constitution. Both are metaphysical relations of sound pedigree. Why should we accept that the one can be involved in giving semantic information its aboutness, but not the other? Such a distinction seems arbitrary. But lacking a principled reason to make this distinc-

tion, it seems like the nominalist's interaction requirement has failed to rule out our gaining semantic information about mathematical objects.

Perhaps the nominalist simply insists that allowing instantiation relations to count as information transferring interaction is a bridge too far. Given that so many ordinary objects enter into constitution relations, it may be granted that such relations can count as a informationally relevant interaction. But it seems like we don't need to admit instantiation as informationally relevant, since the only objects which enter into such relations are metaphysically dubious in the first place. Thus, parsimony compels us to exclude instantiation from counting as interaction. At this late stage, the Weak Gödelian can appeal to the indispensability of mathematics in the scientific enterprise. Specifically, we point out that science becomes unintelligible if we assume that ordinary objects do not enter into instantiation relations with more abstract objects. For example, how would we interpret the statement 'the earth's orbit is an ellipse'? The natural way to do so is to say that the swath of space swept out by the earth as it goes around the sun instantiates the ellipse structure. Positing that ordinary objects enter into instantiation relations has borne great fruit in the sciences. If bearing such fruit is the reason we accept that constitution can count as an informationally relevant interaction, consistency demands that we also admit that instantiation can be informationally relevant.

4.4 Conclusions

So despite its promise, the information-theoretic formulation of Hart's condition has failed to conclusively rule out a Weak Gödelian answer to the access problem. In fact, some of

our considerations would suggest a parallel between gaining semantic information about ordinary objects and gaining semantic information about mathematical objects. In this last section, we briefly speculate on some further defenses and developments which would be a natural extension of the present discussion, before summarizing and concluding.

One important issue that will need to be addressed in later discussions is the exact relationship between semantic and physical information. There are a number of different ways in which this relationship could be more closely examined. First, there is room for discussion about the specific metaphysical relation which holds between semantic information and physical information. The reader will have noted that throughout our discussion, we used phrases like 'embodies' or 'consists in' or 'underlies'. These all go some ways toward capturing how the two are supposed to relate, but it is clearly loose talk. It would make discussion clearer if we could instead identify the relation as some familiar metaphysical relation e.g., supervenience. On a related note, there is still the question of whether physical information and semantic information actually overlap in any interesting way in the first place. We granted that they did as an argumentative concession to the nominalist, but a full treatment would need to address this issue more directly. Finally, at the technical level, we could look at how semantic information relates to other formulations of quantitative information. For example, we might be interested in how Kolmogorov information relates to the semantic conception. The Kolmogorov information (also called Kolmogorov Complexity) contained in a string is equal to the length of the shortest Turing Machine program which produces that string. Importantly, using a different qualitative notion may allow us to bypass the difficulty of specifying which interpretation of probability is relevant to physical and semantic information. Further, comparing different explications of an

informal notion of information may tell us about about how robust this informal notion is.[14]

Suppose for a moment that these investigations could be carried out to a satisfactory conclusion, and the ambiguities present in our current exposition of the Weak Gödelian project dispelled. Even so, we face the daunting task of giving an ontology which can accommodate this project. Such an ontology would have to be fairly radical. After all, in presenting the Weak Gödelian position, we concluded that our sensations can yield perceptions of abstract objects like the 3-structure. Such a position leads us to ask: what is the world like at the fundamental level that this is possilbe? I don't pretend to have a good answer, let alone any arguments for an answer. Still, a complete Weak Gödelian account would be able to situate itself within a broader metaphysical framework.

So the development of the Weak Gödelian project presented here is still at the programmatic level. However, it would be wrong to discount the work we have done to defend the plausibility of a Weak Gödelian answer, which has been considerable. Recall the dialectic: Maddy proposes a Weak Gödelian response to the access problem for mathematical knowledge. However, her commitment to a number-property view raised issues for this response. To get around these issues, we took a number-object approach to the Weak Gödelian project within the context of *ante rem* structuralism. However, our assertion that the objects of mathematical perception are abstract opened us up to the following objection: even if the proposed neurophysical system responsible for mathematical perception is analogous to the system responsible for ordinary object perception, there is a disanalogy between the objects perceived by the one system and those perceived by the

[14]cf. The independent discovery of equivalent formalizations for the notion of computability.

other. However, attempts to spell out how this disanalogy is relevant to perceptual knowledge were inconclusive. We very often found that any criterion which was meant to rule out perception in the abstract case also ruled it out in ordinary cases which the nominalist doesn't want to deny. Given this, we conclude that the purported disanalogy between the objects perceived by the two faculties is not as significant as it was made out to be. As we noted, much more work is needed before the number-object approach to Weak Gödelian epistemology can be considered a complete position. However, our discussion here has shown that the realization of such a position is perhaps more plausible than would have been thought previously. Having established as much, we can end our investigation for the time being.

Bibliography

[1] Amalric, M., Dehaene,S. (2016). Origins of brain networks for advanced mathematics. *Proceedings of the National Academy of Sciences* 113 (18) 4909-4917

[2] Arsalidou, M., Pawliw-Levac, M., Sadeghi, M., Pascual-Leone, J. (2017) Brain areas associated with numbers and calculations in children: Meta-analyses of fMRI studies. *Developmental Cognitive Neuroscience*, In Press.

[3] Bar-Hillel, Y. and R. Carnap, 1953, 'Semantic Information', The British Journal for the Philosophy of Science, 4(14): 147-157.

[4] Benacerraf, Paul, 1965, "What numbers could not be", Philosophical Review, 74: 47-73.

[5] Benacerraf, P. (1973). "Mathematical Truth". *The Journal of Philosophy*, 70(19), 661-679.

[6] Bennett, C.H., 1973, "Logical reversibility of computation", IBM Journal of Research and Development, 17: 525-532.

[7] Burgess, John, 2009, 'Intuitions of Three Kinds in Gdel's Views on the Continuum', in Interpreting Gdel Kennedy, J. (ed.) Cambridge: Cambridge University Press, 2014.

[8] Brillouin, L., 1951, "Maxwell's demon cannot operate: Information and entropy I", Journal of Applied Physics, 22: 334-337.

[9] Chomsky, Noam (1964), *Current Issues in Linguistic Theory*, The Hague: Mouton

[10] Cipolotti L., Butterworth B. & Denes G. (1991) 'A Specific Deficit for Numbers in a Case of Dense Acalculia', Brain, 114, 6, pg. 2619-2637

[11] Clarke-Doane, J. (2016) What is the Benacerraf problem? In F. Pataut (Ed.), *New perspectives on the philosophy of Paul Benacerraf: Truth, objects, infinity.* Dordrecht: Springer.

[12] Cover, T.M. and J.A. Thomas, (2006), *Elements of Information Theory*, 2nd edition, New York: John Wiley & Sons.

[13] Dastjerdi, M., Ozker, M., Foster, B. L., Rangarajan, V., & Parvizi, J. (2013). Numerical processing in the human parietal cortex during experimental and natural conditions. Nature Communications, 4, 2528. http://doi.org/10.1038/ncomms3528

[14] Delgado-Bonal, A. & Martn-Torres, J.(2016), "Human vision is determined based on information theory". Scientific Reports 6, 36038,

[15] Dretske, F., 1981, Knowledge and the Flow of Information, Cambridge, MA: The MIT Press.

[16] Farah, M., 1990. *Visual Agnosia: Disorders of Object Recognition and What They Tell Us About Normal Vision*, Boston: MIT Press.

[17] Field, H. (1989), *Realism, Mathematics & Modality*. Blackwell.

[18] Floridi, L., 2011, *The Philosophy of Information*, Oxford University Press.

[19] Gabor, D., 1964, 'Light and Information', Progress in Optics, 1: 111-153.

[20] Gödel, K. (1964). What is Cantor's continuum problem? Printed in Gdel 1990, pp. 254-270.

[21] Gödel, K. (1990). Collected Works. II: Publications 1938-1974. S. Feferman, J. Dawson, S. Kleene, G. Moore, R. Solovay, and J. van Heijenoort (eds.), Oxford: Oxford University Press.

[22] Hart, W. D. (1977). Review of Mathematical Knowledge. The Journal of Philosophy, 74(2), 118?129.

[23] Jaynes, E.T., (1957), "Information theory and statistical mechanics Phys. Rev.", 106: 620-630

[24] Jones, M. (2015). 'Numbers in Action'. PhD Thesis, University of Bristol.

[25] Jones, G., and Jones, M. (2000) *Information and Coding Theory*. London: Springer.

[26] Kim, J. (1981). The role of perception in a priori knowledge: Some remarks. *Philosophical Studies* 40 (3):339 - 354.

[27] Kraut, R. (1980), Indiscernibility and ontology, *Synthese* 44: 113-135.

[28] Landauer, R., 1961, 'Irreversibility and heat generation in the computing process', IBM Journal of Research and Development, 5: 183-191.

[29] Leng, M., Paseau, A., & Potter, M. (Eds.). (2008) *Mathematical knowledge*. Oxford: Oxford University Press.

[30] Lewis, D. (1986). *On the Plurality of Worlds*. Wiley-Blackwell.

[31] Lewis, D.,(1991). *Parts of Classes*, Oxford: Blackwell Publishers.

[32] Linnebo, Ø. (2006). Epistemological challenges to mathematical platonism. *Philosophical Studies, 129*(3), 545-574.

[33] MacBride, F. (2008). Can Ante Rem structuralism solve the access problem? *Philosophical Quarterly* 58 (230):155-164.

[34] Maddy, P., (1990), *Realism in Mathematics*, Oxford: Oxford University Press.

[35] Maddy, P (1980). Perception and mathematical intuition. *Philosophical Review* 89 (2):163-196.

[36] Mandler G. & Shebo B. (1982), 'Subitizing: An analysis of its component processes?, Journal of Experimental Psychology, 111, (1): 1-22

[37] Maroney, Owen, "Information Processing and Thermodynamic Entropy", The Stanford Encyclopedia of Philosophy (Fall 2009 Edition), Edward N. Zalta (ed.)

[38] Maxwell, J.C., 1867, Letter to P.G. Tait, 11 December 1867, in Life and Scientific Work of Peter Guthrie Tait, C.G.Knott (author), Cambridge: Cambridge University Press, 1911, pp. 213-215.

[39] O'Reilly, R, Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience*. Cambridge: Massachusets; MIT Press.

[40] Potter, M (2007) What is the problem of mathematical knowledge? In Michael Potter, Mary Leng & Alexander Paseau (eds.), *Mathematical Knowledge*.

[41] Resnik, M. (1997). *Mathematics as a Science of Patterns*. New York ;Oxford University Press.

[42] Shapiro, S. (1997). *Philosophy of Mathematics: Structure and Ontology*. Oxford University Press.

[43] Shapiro, Stewart (2011). Epistemology of mathematics: What are the questions? What count as answers? *Philosophical Quarterly* 61 (242):130-150.

[44] Strawson, P. F. (1988). Perception and its objects. In Jonathan Dancy (ed.), *Perceptual Knowledge*. Oxford University Press.

[45] Piazza, M., Pinel, P., Le Bihan, D., Dehaene, S., 2007. A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron* 53 (2), 293-305.

[46] Piazza, M. (2010). "Neurocognitive start-up tools for symbolic number representations". *Trends in Cognitive Science* 14(12): 542-551.

[47] Penrose, O., 1970, Foundations of Statistical Mechanics, Oxford: Pergamon Press.

[48] Roggeman, C., Santens, S., Fias, W., Verguts, T., 2011. Stages of nonsymbolic number processing in occipitoparietal cortex disentangled by fMRI adaptation. J. Neurosci.: Off. J. Soc. Neurosci. 31 (19), 7168?7173.

[49] Santens S, Roggeman C, Fias W, Verguts T (2010) Number processing pathways in human parietal cortex. *Cereb Cortex* 20:77?88.

[50] Steiner, Mark (1975). *Mathematical Knowledge*. Cornell University Press.

[51] Strawson, P. 1966, *The Bounds of Sense*, London: Methuen.

[52] Shannon, C., 1948, "A Mathematical Theory of Communication?, Bell Systems Technical Journal, 27: 279?423, 623?656.

[53] Szilard, L., 1929, 'On the Decrease of Entropy in a Thermodynamic System by the Intervention of Intelligent Beings', Zeitschrift fur Physik 53: 840-856. English translation in The Collected Works of Leo Szilard: Scientific Papers, B.T. Feld and G. Weiss Szilard (eds.), Cambridge, Massachusetts: MIT Press, 1972, pp. 103-129.

[54] Warren, Jared (2017). Epistemology versus non-causal realism. *Synthese* 194 (5).