To appear in the European Journal of Philosophy

Benacerraf's Dilemma Revisited

Bob Hale and Crispin Wright

1. The Dilemma

One of the most influential articles[1] in the last half century of philosophy of mathematics begins by suggesting that accounts of mathematical truth have been motivated by two quite distinct concerns:

(1) the concern for having a homogeneous semantical theory in which the semantics for the statements of mathematics parallel the semantics for the rest of the language

(2) the concern that the account of mathematical truth mesh with a reasonable epistemology [403]

Observing that the two concerns are liable to pull strongly in opposite directions, he proposes two conditions which an acceptable account of mathematical truth should satisfy. One—the semantic constraint—has it that:

> any theory of mathematical truth [should] be in conformity with a general theory of truth … which certifies that the property of sentences that the account calls 'truth' is indeed truth [408]

The other—the epistemological constraint—is that:

> a satisfactory account of mathematical truth … must fit into an over-all account of knowledge in a way that makes it intelligible how we have the mathematical knowledge that we have. An acceptable semantics for mathematics must fit an acceptable epistemology [409]

Whilst some further explanation—of what Benacerraf takes to be required for an acceptable epistemology—is obviously needed before the exact force of the epistemological constraint can be clear, its general drift and point is evident. The precise import of the semantic constraint, however, is less immediately apparent. It is clear that Benacerraf takes it to require, at a minimum, giving a uniform semantic account of our language as a whole, including mathematical language. But sometimes he seems to suggest a stronger requirement. Observing that his examples [405]:

(1) There are at least three large cities older than New York

(2) There are at least three perfect numbers greater than 17

share, on the face of it, the same 'logicogrammatical' form,[2] which he represents as:

(3) There are at least three *FG*s that bear *R* to *a*

he writes:

> … if we are to meet this requirement [the semantic constraint], we shouldn't be satisfied with an account that fails to treat (1) and (2) in parallel fashion, on the model of (3). There may well be <u>differences</u>, but I expect these to emerge at the level of the analysis of the reference of the singular terms and predicates.[408]

Taken on its own, this suggests a very exacting reading of the semantic constraint, under which it can be satisfied <u>only</u> by an account of the truth-conditions of (2) which respects its surface syntax exactly as just described. It is doubtful that Benacerraf can have intended quite such a demanding interpretation. Certainly, so understood, it would go well beyond the somewhat vague demand that 'whatever semantical account we are inclined to give of … singular terms, predicates, and quantifiers in the mother tongue include those parts of [it] we classify as mathematese' [408]. That would seem to come to no more than the weaker requirement—which Benacerraf certainly does endorse—that an account of the truth-conditions of mathematical statements should accord with a broadly referential (i.e. Tarskian) semantics for the language as a whole. This would leave room for accounts of mathematics which view the surface grammatical form of mathematical statements as more or less misleading as to their logical form, provided that their (alleged) logical form involves only devices amenable to Tarskian treatment.

Although Benacerraf does not expressly present himself as posing a dilemma for accounts of mathematical truth, he is widely taken to have done so. Prominent among the accounts of mathematical truth Benacerraf discusses is what he terms the 'standard' or 'platonistic' account, which analyses (2) as being of the form (3)—and, more generally, takes the surface syntax of mathematical statements at face value as an accurate guide to their truth-conditions. This account, Benacerraf argues, complies with the semanticconstraint but is in deep trouble when it comes to the epistemological constraint. A satisfactory general epistemology will, he holds, require, if a subject *X* is to know that *p*, that there should be a suitable causal connection between *p*'s truth-conditions on the one side and *X*'s grounds for believing that *p* on the other. Since, according to the platonistic account, the truth of a true mathematical statement consists in the obtaining of suitable relations among numbers or other abstracta—not located in space or time and so outside the causal swim—it is quite unclear how any such suitable causal relationship could hold, and so quite unclear how mathematical knowledge can be so much as possible. One requirement, then, on a satisfactory account of the semantics of mathematical language is that it should associate mathematical statements with what can properly be taken to be conditions for their truth. Another is that it should not have the consequence that no true mathematical statements can be known to be true. Neither requirement taken by itself appears unsatisfiable. The problem is to give an account which simultaneously satisfies both.[3]

It deserves emphasis that it is a dilemma that Benacerraf is posing. Some commentators give the impression that Benacerraf's primary purpose—or at least the principal effect of his paper—was to present a serious, and perhaps lethal, objection to platonism. This involves, in our view, a quite serious misapprehension. Of course, Benacerraf's argument does raise a problem for platonists, and our principal aim in what follows will be to explain how we think a certain kind of platonism may circumvent it. But the general issue of reconciling semantics and epistemology for mathematics is not just a challenge for would-be platonists—it faces all philosophical positions which allow that pure mathematics presents, as it is normally taken to do, a substantial proper part of human knowledge.

2. Conservative Responses to the Dilemma

The responses to Benacerraf's dilemma canvassed in recent and contemporary work manifest nothing if not variety and resourcefulness. They range, for instance, from Hartry Field's radical view that mathematical statements are not so much as true, let alone known,[4] to Geoffrey Hellman's form of eliminative structuralism[5] which aims to trade-in any appearance that mathematics deals in a worryingly abstract ontology for a systematically modal and hypothetical reading of its statements, to the indispensability argument of Quine: the entities of mathematics are theoretical posits, the justification for believing in which is no different to our justification for believing in the entities posited by theoretical science in general, and is afforded by the indispensability of mathematics within empirical science itself. There is, naturally, much to say about each of these three types of response.[6] But our concern here is with the prospects for the very different—as we shall say, conservative—style of response which takes the Dilemma head-on, unrepentently maintaining both that pure mathematics is correctly construed at syntactic face-value and that, so construed, it represents, at least for the greater part, a body of a priori knowledge.

One kind of conservative view is the sort of platonism sometimes associated with Gödel.[7] On this view mathematics does indeed deal with sui generis kinds of entities, but the impression that these are epistemologically problematic is merely the product of a needlessly restricted—'naturalistic'—conception of the epistemological resources of the human mind., overlooking the accessibility of these entities to 'something like perception' The clear obligation on any such view is to provide an account of what exactly the alleged additional epistemological resources are, overlooked by the broad naturalism which fuels Benacerraf's dilemma, and of how they do indeed enable us to gain knowledge of the entities that compose the special subject matter of mathematical theories.

Two broad approaches seem possible: intuitional and intellectual. It may be proposed, first, that an epistemology of mathematics should reckon with a special faculty—traditionally 'intuition'—which enables an awareness of systems of abstract, and especially mathematical objects and of their characteristic properties, broadly as ordinary sense perception makes us aware of ordinary concrete objects and their properties. Or it may be proposed that access to the objects of pure mathematics is afforded by our general abilities of reason and understanding. We shall consider examples of each approach in turn.

It's fair to say that the former kind of approach is now widely regarded as extremely unpromising. The immediate and principal difficulty with all 'special faculty' proposals is that while they come cheap—are readily available to anyone who wants to take a broadly realist view of any subject matter, for instance ethical or aesthetic value, where the claim of objectivity is prima facie in tension with a naturalistic outlook—they are also very difficult to discipline sufficiently to offset the charge of superstition; and the paradigms that we have of such discipline, for instance the kind of empirical controls we would place on the experimental corroboration of such phenomena as telepathy or precognition, apply only in circumstances where there is independent corroboration of how matters stand with the subject matter to which the faculty in question is allegedly responsive. That factor goes missing in the present case, for whether there are indeed the object-involving mathematical states of affairs to which the putative intuitive faculty would give access is a matter which has to be regarded as sub judice along with the faculty itself. Nevertheless at least one distinguished philosopher of mathematics has, for some decades, conducted a patient and sympathetic exploration of the prospects for explicating the notion of mathematical intuition, construed on the perceptual model as a faculty of awareness of abstract objects and, thereby, providing a source knowledge of truths about them. We shall, all too briefly in the space available to us, rehearse some grounds for pessimism about whether the concept of mathematical intuition that has been emerging from the writings of Charles Parsons is likely to service an even locally successful conservative response to Benacerraf's dilemma.

3. Parsons on Intuition

To simplify matters, let us focus on the case of number theory. Central to Parsons' account of the matter is a distinction between intuition of objects and what may be called propositional intuition or intuition that. Since the analogy between (mathematical) intuition and sense-perception Parsons is anxious to conserve[8] is to consist, roughly, in the idea that much as perceptual knowledge that, say, there is a duck on the pond depends upon and is in some way furnished by perception of the duck, so intuitive knowledge that, say, 7+5=12 will in some way be based upon intuition of some appropriate objects, it is to the notion of intuition of objects that he wishes to assign a fundamental rôle. The example might suggest that Parsons would take its objects of intuition to include (at least small) natural numbers, but in fact, that is not his view. He draws an important distinction between quasi-concrete and pure abstract objects. In extension, at least, this coincides with the distinction among abstract objects between those which closely depend upon—which 'belong to'— concrete objects of one or another kind, and those which do not.[9] Linguistic expressions, considered as types as distinct from tokens, constitute a clear—and, for Parsons, particularly important—example of quasi-concrete objects, while pure sets and numbers of various kinds are plausibly taken to stand in no such dependence relation to concrete objects. In terms of this distinction, the objects of mathematical intuition, as Parsons conceives it, are restricted to the quasi-concrete. In particular, he denies that there can be intuition of the natural numbers themselves, and would presumably take the same view in regard to any other type of pure abstract objects, such as the pure sets.

Parsons' explication of intuition of is heavily dependent on one particular example, based on a 'language' whose sole primitive symbol is the stroke '|' and whose well-formed expressions are arbitrary strings of strokes: |, ||, |||, …. As he says, this sequence of stroke-strings is isomorphic to the natural numbers, if one takes the singleton string as 0 and the operation of adding one more stroke on the right of a given string as the successor operation.[10] Ordinary perception of a string of strokes will be perception of a concrete inscription, i.e. of what is usually called a token, as opposed to a type. In this kind of case, the objects of intuition are stroke-string types. Although one may properly speak of hearing or seeing words or sentences, meaning types,[11] a perception of a stroke-string token is not as such an intuition of a stroke-string type:

> One has to approach it [i.e. the object of perception] with the concept of a type, first of all to have the capacity to recognise other tokens as of the same type or not. Something more than the mere capacity is involved, which might be described as seeing something as the type.[12]

To a first approximation, we might say that intuition of an object, for Parsons, is just what occurs when we (literally) perceive something and, equipped with an [appropriate] concept of type, 'see it as' a certain type. Thus relatively straightforward perceptual recognition of expression-types counts, for him, as a kind of intuition of objects.[13] Full-blown intuition of types, involving conscious deployment of more sophisticated conceptual resources, may be seen—or so we take him to be suggesting—as continuous with such cases, and not different in kind from them. Certainly he is keen to

> … dispel the widespread impression that mathematical intuition is a special faculty, which perhaps comes into play only in doing pure mathematics. At least one type of essentially mathematical intuition, of symbol- and expression-types, is perfectly ordinary and recognised as such by ordinary language. If a positive account of mathematical intuition is to get anywhere, it has to make clear …. that mathematical intuition is not an isolated epistemological concept, to be applied only to pure mathematics, but must be so closely related to the concepts by which we describe perception and our knowledge of the physical world that the 'faculty' involved will be seen to be at work when one is not consciously doing mathematics.[14]

A concern arises here about whether the conception of intuition being proposed doesn't reduce to an unwarranted inflation of ordinary intelligent perception. Why think that there is some extra faculty with its own special objects involved in such cases, rather than merely concrete perception and the exercise of concepts? Parsons is perhaps trying to forestall this worry when he writes, a little further on in his (1980),

> … in all these cases it seems not to violate ordinary language to talk of perception of the universal as an object, where an instance of it is present. This is not just an overblown way of talking of perceiving an instance as an instance … because the identification of the universal can be firmer and more explicit than the identification of the object that is an instance of it.[15]

One may wonder whether the appeal to consonance with ordinary language can be sufficient to address the concern, or what the claimed relative firmness of the 'identification of the universal' amounts to.

Setting that issue to one side, it is anyway clear that intuition of objects of a certain kind is of no use unless it is somehow at the service of propositional knowledge about those objects. As a simple example, Parsons gives the statement that ||| is the successor of ||. In virtue of the fact that a perception or imagining of certain tokens of the relevant types can 'play a paradigmatic role', this statement, he says, can be seen to be true on the basis of a single intuition. But that is a relatively easy case. More interestingly, but more controversially, he claims that intuition can give us knowledge of analogues for strings of the elementary Peano Axioms:[16]

(PA1′) | is a stroke string

(PA2′) | is not the successor of any stroke string

(PA3′) Every stroke string has a successor which is also a stroke string

(PA4′) Different strokes strings have different successors

The problem, of course, is to see how, following Parson's intuitive route, knowledge of <u>general</u> truths about intuited objects, like (PA2′-4′), can be achieved. We may focus on

(PA3′), which Parsons takes to be equivalent to the assertion that each string of strokes <u>can</u> be extended by one more, and to constitute 'the weakest expression of the idea that our 'language' is potentially infinite'.[17] He agrees that it cannot be intuitively known via mathematical intuition founded on actual perception. But we can solve the problem, he claims, by extending the notion of intuition so as to allow that intuition of a type may be founded on <u>imagination</u>, as distinct from perception, of a suitable token:

> But if we imagine any string of strokes, it is immediately apparent that a new stroke can be added. One might imagine the string as a <u>Gestalt</u>, present all at once: then, since it is a figure with a surrounding ground, there is space for an additional stroke … Alternatively, we can think of the string as constructed step by step, so that the essential element is now succession in <u>time</u>, and what is then evident is that at any stage one can take a further step.[18]

As Parsons emphasises, the crucial thing—if the deliverance of intuition is to have the requisite generality—is that one has to imagine <u>an arbitrary string of strokes</u>, and that seems to leave us facing a problem of much the same sort as besets Locke's abstract general triangle. We may evade it, he suggests, in one of two ways—by

> imagining <u>vaguely</u>, that is imagining a string of strokes without imagining its internal structure clearly enough so that one is imagining a string of <u>n</u> strokes for some particular <u>n</u>, or taking as paradigm a string (which now might be perceived rather than imagined) of a particular number of strokes, in which case one must be able to see the irrelevance of this internal structure, so that in fact it plays the same role as the vague imagining.[19]

It may be doubted, however, whether either suggestion accomplishes much to the purpose. It is not clear how imagining a vaguely imagined string extended by a further stroke is supposed to furnish us with knowledge that <u>any</u> string can be extended, rather than merely that <u>some string or other</u> can be. And as for the second suggestion, what obviously does all the work is the idea that we somehow see the irrelevance of the internal structure of the imagined (or perceived) string. But how is that 'seeing' supposed to work? The idea is clear enough when, in proving a geometrical theorem—Pythagoras' theorem, say—we work with a (rough) diagram of a particular right-angled triangle: seeing the irrelevance of the specific values of the other angles, and the specific lengths of the sides, then consists in the fact that none of the steps in our accompanying reasoning relies on any assumptions about them. But—precisely because there is no counterpart to such reasoning in Parsons' supposed intuitive verification of (PA3′)—it remains obscure at best how the idea that we see the irrelevance of internal structure is to be cashed in.

We shall return to the issue of the alleged intuitive verifiability of (PA3′) below. There is, however, a more general misgiving. Even if Parsons' case as so far presented could be sustained, there would clearly be significant limitations upon the extent to which intuition and intuitive knowledge would have been disclosed as contributing to an account of arithmetical knowledge. An intuitively founded recognition that (PA1′- 4′) are true of stroke-strings, conceived as types, would amount to knowledge of a model of the Dedekind-Peano axioms. But of course knowing that there is a model of the Dedekind-Peano axioms doesn't amount to knowledge of their truth. You could know of some abstract system that it is a model of an axiomatisation of Newtonian mechanics, but that obviously doesn't amount to knowledge that Newtonian mechanics is true. So Parson's account has left us short of what we wanted: an account of our knowledge that the Dedekind-Peano axioms (or even just the elementary ones) are true. How might the gap be closed?

Well, it is difficult to see how one could get to the corresponding Dedekind-Peano axioms by any sort of generalisation from (PA1′-4′). The latter concern objects—albeit (impure) abstract ones—of a quite specific sort, which not only have their roots, so to speak, in concrete objects of a certain kind (stroke inscriptions), but are themselves in a sense spatial. While the Dedekind-Peano axioms likewise, on a face-value construal, concern a definite collection of objects—the (pure) natural numbers—it is of their essence that statements about them have, as Frege stressed, application to anything thinkable, including the non-spatial The difficulty is therefore to see how, if the four relevant axioms could somehow be inferred from their stroke-string analogues, they could enjoy that essential generality (indeed, universality).

And even prescinding from *that* problem in turn, there is anyway an obvious question to be faced by any Parsons-style intuitionist about full knowledge of arithmetic. Parsons is always careful to stress that he thinks intuition can only be *part* of the story, if only because no general principle of induction is intuitively knowable. How, even if intuitively founded knowledge of the four elementary axioms might somehow be secured, is the rest of the epistemological story to run, and how is it to make contact with the (in certain respects) quite modest claims Parsons does wish to make for intuition?

We can envisage one quite straightforward proposal, which might seem attractive as providing a way around the problem of generality noted above. This would be to combine the claim that intuition can deliver knowledge of a model of arithmetic with some species of *eliminative structuralist* account—perhaps a modalised version along the lines advocated by Hellman—of the content of purely arithmetical statements.[20] On such a view, the generality of arithmetic is to be accounted for by construing its theorems, not as singular statements about a distinctive range of special, abstract objects—the natural numbers—but as general claims about arbitrary ω-sequences. One well-known difficulty for any simple, non-modal, version of this position is that, unless there actually exists at least one such sequence, its re-interpretations of arithmetic falsehoods will (vacuously) come out true along with the truths. A partial solution—advocated by Hellman, following Putnam, and noticed by others—is to adopt a modalised translation scheme, under which each sentence $\underline{A}$ of arithmetic is rendered, not as a universally closed material conditional:[21]

$$\ni\underline{s}\ (\underline{s} \text{ is an } \omega\text{-sequence} \rightarrow \underline{A}'(\underline{s}))$$

but as its necessitation:

$$[\_]\ni\underline{s}\ (\underline{s} \text{ is an } \omega\text{-sequence} \rightarrow \underline{A}'(\underline{s}))$$

The vacuity problem can then be avoided, without postulating the existence of any actual ω-sequence, provided that it is at least possible that there should exist such a sequence.

This solution does not come for free,[22] however, since there remains an epistemological debt to discharge. The required possibility claim, though weaker than the claim that there actually exists an ω-sequence, is—like most philosophically significant possibility claims—far from negligible, or such as to command assent by default. How is it to be justified? Here, it may seem, is precisely where Parsons' limited but significant claims for intuition might play a crucial rôle. Why can't we bridge the epistemological gap quite simply, by appealing to our intuitively grounded knowledge of the infinity of stroke-strings, as obtainable by reasoning from (PA1′-4′), and thereby enforcing the required possibility claim?

We are sceptical. Recall, first, that we are distinguishing between token strings and types. Theinfinity enjoined by (PA1′-4′) is that of the types, not the tokens. So now we have to rescrutinise Parsons' claim that (PA1′-4′) may be intuitively known. The crucial question is: what are the conditions for the existence of types, e.g. of type strings?

There seem to be three positions one could take. The strongest requirement would be that a type string exists (if and) only if there actually is at least one token of that type (i.e. types must be instantiated). In that case, there will be infinitely many type strings only if there are infinitely many tokens. More generally, knowing there to be infinitely many quasi-concrete objects, under this assumption about type-existence-conditions, would require knowledge of the actual infinity of some collection of concrete objects. Obviously this is no good. We do not have any such knowledge. And, in any case, the supposition of such a conrete infinity seems quite foreign to any plausible epistemology of arithmetic.

At the other extreme, one might insist that the existence of types is entirely independent of their actual or even possible concrete or imaginative instantiation. But this is a rampant platonism and whatever other flak it may draw, it seems clear that Parsons could not defend the claim without allowing that types are after all not quasi-concrete objects at all, but pure ones. That would be fatal to the intuitional route—at least as Parsons conceives it—since he denies that one can have intuitions of pure abstract objects.[23]

There is a third, final possibility: that a type string exists (if and) only if there could be a token of that type (i.e. there may be uninstantiated types, but only ones which could be instantiated). On this proposal, making out the infinity of the string types will require making good the claim that there is at least a potential infinity of concrete inscriptions of a certain sort, i.e. that any sequence of concrete inscriptions could be extended. Here the claim of potential infinity needs to be taken seriously. It would be unfair to the proposal to insist that what is required is the possibility of an actual infinity of concrete inscriptions, i.e. one world in which infinitely many concrete tokens exist (and in which physical space is accordingly infinite, for instance.) Again: such a possibility claim is not obviously even true, and is in any case quite foreign to any plausible epistemology of arithmetic. But the supporter of Parsons has no need to make it. The claim should be, rather, that for each n, an n-fold concrete stroke-string is possible. And that is believable enough.

Believable enough, but the question is what the supporter of Parsons may offer as its epistemological ground. That ground must, it seems, —in keeping with Parsons' general framework—to be the kind of imagination which when actual tokens give out, so that their perception can no longer be the ground of the intuition of types, is supposed to step into the gap to provide imaginary tokens to perform the same role. So the claim has to be that given any perceived or imagined token stroke-string, a single-stroke extension of it is imaginable. And while there may be a sense in which that is true, it seems to us that the claim cannot carry the weight placed on it on the present context.

The problem is: what is the modality implicit in 'imaginable'? In particular, does it involve rigidification of our actual imaginative capacities or does it not? If it does, then the claim involves that, for each imagined stroke-string token, we are actually capable of going the extra millimetre, so to speak, and accomplishing an imaginative representation of a successor. There doesn't seem to be the slightest reason to believe that claim—our actual imaginative capacities are presumably no more infinite than our vision, or intelligence, or any other of our cognitive and intellectual abilities. And once again, even if that is doubted, the supposition to the contrary seems quite foreign to a plausible epistemology of arithmetic. So the implicit modality must be flexible: we are allowed to consider extensions of our actual imaginative powers. But now the question is, what disciplines admissible such extensions?—what determines what would be imaginable if our powers of imagination were extended? And the present authors cannot see what other answer could be returned but (for the particular case) that this is determined by the range of forms—types—that there are for token stroke-strings to assume! The dilemma, in other words, is that while imaginability in practice, by beings as we actually are, is too impoverished a basis to secure a basis for an infinity of stroke-string types in line with the third proposal, it is at best obscure how to liberalise the notion so as to get the desired result without invoking an anterior conception of the range of types which concrete tokens—perceived or imagined—might in principle exemplify. And now the grounds for that anterior conception cannot, it appears, be located in intuition when that is conceived in Parsons way, as based on the inputs of perception and imagination.

While we cannot attempt to argue further for the point here, we do not think this difficulty is special to Parsons' particular approach. The point is general, afflicting any attempt to ground our knowledge of the truth of infinitary mathematical axioms in our apprehension of possibilities supposedly grounded in imagination. This is a motive for thinking that the real source of our knowledge of even simple infinities must lie on the other side of the earlier divide: that it is delivered not by intuition, so long as that is viewed on a broadly perceptual model, as a mode of cognition of objects underlying propositional knowledge about them, but by the operations of intellection and understanding. The second broadly conservative strategy is to attempt so to construe our access to mathematical objects in general. Our own preferred neo-Fregean approach belongs firmly in this camp. But it is a broad church, encompassing also Stewart Shapiro's particular brand of ante rem structuralism.[24] In the next section, we briefly review Shapiro's approach and highlight our most serious misgiving about it.

4. Intellectualist responses to the Dilemma (i): Shapiro's structuralism

For Shapiro, pure mathematics is concerned with structures conceived as self-standing abstract entities. Arithmetic, for instance, is concerned with the structure exemplified both by the decimal numerals and by Parsons' system of strokes and the natural numbers themselves are to be viewed as the places in this structure. It would be natural to think in the first instance of talk of structures of mathematics as talk about properties — about ways particular collections of objects may be ordered and organised. The distinctive idea of the ante rem structuralist is so to reify these properties that the resulting complex objects are themselves instances of them. Thus the structure: ω-series, consists of an array of places so ordered as itself to consitute an ω-series. The Platonic resonances of the idea are salient.

Since he so conceives the structures that mathematics studies, the ante rem structuralist squarely confronts Benacerraf's dilemma. Structures, so conceived, are timeless, abstract entities. How do we make their acquaintance, and how do we recognise what is true of them? Shapiro makes three proposals. Small finite structures, he suggests, are accessible to us via processes of abstraction-from-differences and pattern recognition. It is thereby, Shapiro believes, that we advance to conceptions of, for example, the types illustrated by inscription tokens and the notes exemplified by a particular token sounds; and Shapiro views the same kinds of processes as leading to a grasp of, for instance, 'the *four pattern* .. the structure common to all collections of four objects'.[25] Obviously this way of 'accessing' structures depends on confrontation with instances of them, and issues therefore are immediately raised concerning our access to simple infinite structures, like ω -series where there can be no confrontation with a (completed) instance. Indeed an issue arises earlier concerning our grasp of large finite structures where, again, no humanly recognisable pattern is exemplified by the instances. Here, if we understand him correctly, Shapiro proposes that the essential step is that the learner advances to a conception of a second order pattern — a pattern exemplified, now, by the types grasped by a reflection on concrete tokens of them. For instance, the key to grasping the idea of stroke sequences too long to be surveyably written down, and indeed of the infinity of the series of such sequences, is for the learner to grasp the pattern exemplified by the types which are in turn exemplified by the tokens in the following sequence:

| || ||| |||| ||||| |||||| [26]

The reader may feel some uneasiness. Even granting Shapiro that we are capable of accessing simple finite structures as objects by perceptual confrontation with concrete tokens of them, is it supposed to be just obvious that we should then be capable of, so to say, ascending a level and contemplating arrays of such structures and, thereby, accessing higher-order structures as objects in their own right? Is it just obvious that we have any ability of which this, to anyone but a committed ante rem structuralist, would seem to be the correct description?

Shapiro, though, betrays no uneasiness on this point. What he does recognise, of course, is that the means described provide access to at most quite small structures — from the standpoint of contemporary mathematics and set theory — and his discussion moves to the question of the possible means whereby more elaborate structures can be brought to our attention. To this purpose, he proposes two, each conceived as modes of implicit definition. The first is something akin to Fregean abstraction: abstraction, that is, not as a process which forms a concept by prescinding from differences among instances of it, but in the now familiar sense whereby one forms a conception of the identity conditions of a new kind of object in terms of an equivalence relation on objects of a familiar kind. (Shapiro does not say so, but the abstraction of inscription-types from inscription-tokens may be viewed as involving involving a move of this sort.)[27]

Abstraction in this sense — Fregean abstraction — will occupy us in due course and is indeed, in our view, a crucial resource in coming to understand how the objects of mathematics can be intelligible to us, and how mathematics may knowledgeably investigate them. But this is because it sustains a crucial contrast, in our opinion, both to Shapiro's first method — (extended) pattern recognition — and to his third.

The third method is breathtakingly simple: it is simply one of characterisation. In the best case, Shapiro holds, by giving a (categorical) characterisation of an intended structure, we thereby communicate the structure — make it available as an object of intellection. And once so made available, it may be investigated by exploring the deductive and model-theoretic consequences of the characterisation by which it was communicated.

This third method, if allowable, naturally trumps the others. No need, for instance, to fool around with pattern recognition among inscription tokens, and pattern recognition among the types thereby (putatively) accessed — the second order (categorical) Dedekind-Peano axioms communicate everything necessary to grasp the structure: ω-series. Sure, there may be a role for more basic intellectual operations in the story of how we can arrive at an understanding of the language in which the Dedekind-Peano axioms are stated. But one would suppose that that additional detail would belong to empirical psychology, rather than philosophy. If Shapiro is right, all he needs to do in order to explain how the structure, ω-series, is accessible to us as an object of mathematical investigation is to call attention to the fact that we are capable of grasping a canonical axiomatic description of it.

When does a description communicate — make accessible as an object of intellection — an <u>ante</u> <u>rem</u> structure? Shapiro does not want to allow that just any old consistent description so serves. Tighter constraints are wanted, and Shapiro flags them in his notion of <u>coherence</u> — where roughly, a characterisation is coherent just in case it is satisfiable in the standard hierarchy of sets. (That, anyway, is the intended <u>extension</u> of the notion: it would be a serious concern if the best that can be done to explicate it is simply to help oneself, in that way, to an assumed prior ontology and epistemology of sets.) There is much to say about coherence in this setting and we shall not attempt to respond to Shapiro's interesting and resourceful discussion of it here. Our principal reservation with his approach is not, however, that he can give no satisfactory account of the restrictions that need to be imposed before an axiomatic description should be regarded as characterising a structure in the sense he intends.. It rather concerns the very idea that, merely by giving such a description, we can do more than <u>convey a concept</u> — that we can, in addition, induce awareness of an articulate, archetypal <u>object</u>, at once representing the concept in question and embodying an illustration of it. Someone who writes a fiction, even the most coherent fiction, does not thereby create a range of entities whose properties and relations are exactly as the fiction depicts. Rather— it can be agreed on all hands— she creates a concept, a description of a possible scenario in which certain things, real or imagined, might be so qualified and related. It is implicit in Shapiro's view, by contrast, that there can be no such thing as a fictionalised structure. Try to write about merely imaginary structures, as about imaginary people, and the very description of your fiction, if coherent, will defeat your purpose. Willy-nilly, a Platonic entity will step forward to fulfil your descriptive demand.

Shapiro is, naturally, quite self-conscious about this aspect of his view. If he's right, mathematical fictionalism is simply an incoherent philosophy of mathematics from the start. With structures, coherence suffices for existence. But that would be a very momentous contention, and we have been unable to find any compelling argument for it in his discussion. Once granted, it provides for the simplest possible conservative response to Benacerraf's;s Dilemma: a coherent characterisation at once communicates the nature of a structure and gives an example of it; and this example may then be investigated a priori by deduction from the characterisation, with the results guaranteed to hold of instances of the relevant structure in general. Against this, we wish to set the orthodox idea that in mathematics, as elsewhere, there is a gap between concept and object, that it is one thing to give the most precise characterisation of a way an object, or field of objects, might be and another to have reason for thinking that there actually are some objects which are that way. Shapiro's strategy is, we believe, correct: in order for abstract objects to be available to intellection, we have somehow to pass from the implicit definition of a concept to a recognition that it is instantiated. But the passage cannot be that direct.

5. Intellectualist responses to the Dilemma (ii): Neo-Fregean Platonism

Why have philosophers been prone to find the idea troublesome that we may have knowledge about abstract objects? There is no doubt that broadly causal conceptions of knowledge of the kind Benacerraf gestured at in stating his dilemma, and more generally, sympathy with reliabilist and naturalistic trends in epistemology, have been a major factor. But a sense of vertigo in the face of abstracta has, we think, other and more general sources, albeit ones that are less sharply defined and seldom unmistakably articulated.

One is the tendency to operate with a conception of abstract objects—as 'outside' space and time—which is not merely metaphorical but wholly negative. The metaphor need not, in itself, be particularly damaging, so long as we remind ourselves, when necessary, of its literal content: roughly, that—at least in the general run of cases—it makes no sense to ask where an abstract object is, or when it came into existence, or how long it will last. It is, rather, the negative aspect of the characterisation that does damage. If we focus exclusively on what abstract objects are not, with no thought about what they are or might be supposed to be, we can scarcely expect much of a sense of direction when we try to consider how we might get to know about them.

A second is the idea that knowledge of truths about objects of any kind must involve some form of prior interaction or engagement with those objects. That notion is naturally taken to call for some sort of physical connection—perhaps of the sort that occurs in normal sense perception, or at least some less direct form of physical influence upon us—and so is obviously inimical to the abstract. But it is not obviously a notion that must be accepted. To be sure: if we forget the 'prior' and the notion is given a sufficiently broad construal, so that possession of any sort of identifying knowledge of an object suffices for 'engagement', the idea reduces, near enough, to a truism—one can hardly be credited with knowledge of truths about objects unless one knows which objects are in question. But so construed, it need raise no hurdle for platonism. The crucial thought—we should say: 'mistake'—is the additional idea that such 'engagement' is presupposed by and must be already in place before any knowledge of truths about objects can be had.

Once locked into thinking about the access problem within this straitjacket, we can hardly avoid a further thought—that the problem is not just how we can know anything about abstract objects, but how can we even so much as think about them at all. For what engagement with a range of objects seems needed for, if anything, is the direction of attention presupposed in thought about them, a fortiori in thoughts involved in their investigation and the acquisition of knowledge about them. So much the worse, it may therefore be said,[28] for platonism, which is therefore in trouble on two counts, not just one: it obstructs both a satisfactory epistemology and a workable theory of reference. Even if one could give a platonist account of the truth conditions for mathematical statements (or any other class of statements supposedly about abstract objects), the objector will say, we would be unable to explain how such statements, so understood, could be known or reasonably believed; but in fact—the more basic objection is—one cannot even give such a semantical account, since one cannot so much as intelligibly articulate platonistic truth-conditions. The fundamental problem is not how, given that mathematical statements are about abstract objects, we could know them to be true, but how they could intelligibly be about such objects in the first place.

That is therefore the question the platonist should address first. The essential steps towards the answer we propose are adumbrated in the central sections of Frege's Grundlagen.[29] At §62, Frege explicitly confronts the access problem—as a problem about how we can refer, in thought or speech, to abstract objects—in what is, for him and for us, the crucial case: 'How, then, is a number to be given to us, if we can have no idea or intuition of it?' His answer begins with, and is shaped by, the famous context principle: 'Only in the context of a proposition does a word mean anything'. From this he infers, first, that what is needed, in order to respond to the access problem, is to 'explain the sense of a proposition in which a number word'—numerical singular term—'occurs' and then, since 'number words are to be understood as standing for self-subsistent objects', that what, more specifically, must be explained—by giving, without using number words, their truth-conditions—is the sense of identity statements connecting terms for numbers. Since, as Frege has already argued, cardinal number is a property of concepts and numbers, correspondingly, are essentially things which belong to concepts, such identity statements will, in the fundamental case, be of the form: The number of Fs = the number of Gs.

The significance of these simple but revolutionary steps can scarcely be over-emphasised: if they are well-conceived, they contain the essential ingredient—or, more accurately, one of the essential ingredients—of the antidote to aporia about abstract reference which platonism needs. For they enjoin that in order to establish an intelligible use for singular terms purporting reference to numbers, or other abstract objects—that is, objects which are not 'external' (located in space), and of which we can have no 'idea' or 'intuition', but which are, in Frege's view, nonetheless objective—it suffices merely to explain the truth-conditions of statements incorporating such terms. No precondition involving prior engagement with or attention to the referents of such terms is soundly imposed. Moreover if, under a suitable explanation of the truth-conditions of an appropriate range of such statements, suitable such statements[30] are—or may warrantedly be claimed to be—true, then those of their ingredient terms which purport reference to numbers or other abstract objects will in fact refer—or may warrantedly be claimed to succeed in referring—to such objects; and the intelligent contemplation of such a statement will constitute thought directed upon the objects concerned.

It is vital to appreciate the scope of the point which Frege was making. The context principle is not just—as is sometimes uncharitably supposed—a convenient device he reaches for as a means of evading special difficulties about talk of abstract objects. If the basic idea is right at all, it is completely general in application, applying to the meaning and reference of words for concrete objects just as much as to talk of abstracta. Singular thought, and object-directed thought in general is, on this view, enabled by and fully realised in an understanding of suitable kinds of statement. It is not something which, in however rudimentary a form, must be engaged in before such statements can be understood, as part of the process of coming to understand them. The opposite idea is precisely what is embodied in the Augustinian conception of language put up for rebuttal at the very outset of the Philosophical Investigations; and the prime spur towards the 'naturalist' tendency which finds abstract objects per se problematical is the idea, at the heart of the Augustinian conception, that some, however primitive, form of conscious acquaintance—and hence, when acquaintance is naturalistically construed, some kind of causal relationship—must lie at the roots of all intelligible thought of, and hence reference to objects of a particular kind. While this is not the place to enlarge upon the relevant points in detail, it is by no means an unusual reading of *Philosophical Investigations* to believe that the book as a whole accomplishes a compelling critique of this idea.[31] For Wittgenstein, language is not a mere medium for the expression of thought but—at least for thoughts of the level of sophistication of pure mathematics—is an activity in which thought has its very being. And singular thought—at that level of sophistication—is consequently not something which underwrites a grasp its expression but is constituted in the understanding of its expression. The philosophical question of how that understanding may be acquired or explained, is therefore something we should approach with an open mind, unfettered by incoherent prejudices about ostensive definition or the mythology, in which nominalism is rooted, of primitive forms of intellectual attention without which the mind cannot engage the real world. The outstanding question, after Frege's re-orientation of the issues, is how an understanding may be acquired of those forms of statement which, if true and if taken to involve the reference to and quantification over abstracta which they seem to involve, serve to give us the means to think of and refer to such objects.

Frege takes the question head-on. The sections of *Grundlagen* immediately following §62 advance the proposal that we explain the truth-conditions of numerical identity statements by means of what has come to be called:[32]

Hume's Principle:

The number of Fs = the number of Gs ↔
there is a one-one correspondence between the Fs and the Gs

In response to the objection that he has no business (re)defining identity for the special case of number, Frege explains that, on the contrary, his aim is 'to construct the content of a judgement which can be taken as an identity, each side of which is a number', so that rather than defining identity specially for this case, he is 'using the concept of identity, taken as already understood, as a means for arriving at that which is to be regarded as identical'. Switching, for expository purposes, to the simpler (but in his view, relevantly similar) case of an explanation of the concept of direction via the

Direction Equivalence: The direction of line a = the direction of line b ↔
lines a and b are parallel

he explains in greater detail:

The judgement 'line a is parallel to line b', or, using symbols,

> a//b
>
> can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'the direction of line a is identical with the direction of line b'. Thus we replace the symbol // by the more generic symbol =, through removing what is specific in the content of the former and dividing it between a and b. We carve up the content in a way different from the original way, and this yields us a new concept.[33]

Frege's talk of carving up the same content in a new way is, of course, a metaphor—albeit a highly suggestive and potentially helpful one—and it is therefore desirable to try to spell out in plainer terms what process of conceptual surgery he may reasonably be supposed to have had in mind. That will be to explain in what way or sense instances of the left and right sides of the Direction Equivalence (or Hume's principle, or, more generally, any similar abstraction principle) may be held to share the same content. As a matter of Fregean exegesis, this task is complicated by at least two historical facts. One is that Frege's use of the term 'content' in Grundlagen and earlier writings is anything but uniform—he speaks both of the content of whole sentences and of the content of subsentential expressions, including names; sometimes by the content of a symbol he seems to mean what it stands for, but he also speaks of conceptual content in ways that suggest something more like what he later calls sense. The other is that soon after Grundlagen,[34] he expressly abandons the notion of judgeable content, replacing it by the contrasted notions of thought and truth-value, as a special case of his general distinction between sense and reference. Since we lack space to pursue this issue in the detail it deserves,[35] we shall simply state, without supporting argument, how we think Frege is best understood.

We may view the Direction Equivalence as an implicit definition of the direction operator—'the direction of …'—and thereby of the sortal concept of direction. The import of the stipulation of the equivalence is simply that corresponding instances of the left and right sides—matching sentences of the shapes 'the direction of line a = the direction of line b' and 'lines a and b are parallel'—are to be alike in truth-value, i.e. materially equivalent. But because the stipulation is put forward as an explanation, its effect is to confer upon statements of direction-identity the *same truth-conditions* as those of corresponding statements of line-parallelism.[36] Thus what a recipient of the explanation immediately learns is that whatever suffices for the truth of a statement of line-parallelism is equally sufficient for the truth of the corresponding statement of direction-identity. However, she also understands that she is to take the surface syntax of direction-identity statements at face value. She already possesses the general concept of identity, and so is able to recognise that the expressions flanking the identity sign must be singular terms. Further, she already understands terms for lines, and so can recognise that 'the direction of …' must be being introduced as a functional expression, denoting a function from lines to objects. From this, she is able to gather that the objects in question simply are objects for whose identity it is necessary and sufficient just that the relevant lines be parallel. She learns, in other words, that directions just are objects with exactly those identity-conditions, and thus acquires the concept of *direction*—it is the sortal concept under which fall any and all objects having those identity-conditions.

On this account, the force of Frege's metaphor is that a statement of direction-identity may be viewed as involving a *reconceptualisation* of the parallelism of a pair of lines. The situation of two lines bearing to one another the relation of parallelism may be conceived as constituting an identity among objects of a 'new' kind—the directions of those lines—rather as, to adapt to this context a rather different example of Frege's own, a single pack of cards may be conceived as precisely that, or as four suits, or as fifty two individual playing cards.[37] In like manner Hume's Principle may be seen as implicitly defining a concept of cardinal number in such a way that statements of numerical identity involve reconceptualisations of one-one correspondences between concepts. And, as Frege saw,[38] the procedure has potentially a much wider application. The Direction Equivalence and Hume's Principle are two instances of a general type of *abstraction principle*—that is, a principle of the shape:

(Abs) $\ni\alpha\ni\beta\ (\Re(\alpha) = \Re(\beta) \leftrightarrow \alpha \approx \beta)$

where $\approx$ is an equivalence relation on entities of the type of α,β, and $\Re$ is a function from entities of that type to objects. Given any suitable equivalence relation, we may abstract over it to introduce an abstract sortal concept—a concept under which fall abstract objects of a certain type, the identity or distinctness of which is constituted by the obtaining of that equivalence relation among entities in its field.

There is a great deal more to be said in explanation, qualification and defence of this procedure of implicit definition by Fregean abstraction. However, it should already be evident that, if it constitutes a viable means of introducing abstract sortal concepts, such as that of <u>direction</u>, and establishing a use for terms potentially denoting their instances, it furnishes also at least a partial means of resolving the problem of epistemological access standardly taken to beset platonism. The problem, in general terms, was to see how statements of a given kind can be understood as involving reference to abstract objects and can yet remain, at least in principle, humanly knowable, given that the objects they concern are outside space and time and in consequence can stand in no sort of epistemologically relevant, causal relations to human knowers. And the solution—in the case of directions—is that, provided that certain kinds of statement about lines can be known to be true, there can be no further problem about our capacity for knowledge of the truth of members of a certain basic class of statements about directions. For if the concept of direction is introduced by Fregean abstraction, statements of direction-identity have precisely the same truth-conditions—are true, if true at all, in virtue of the very same states of affairs—as statements asserting relevant lines to be parallel. And so it is with a basic class of statements about numbers. A statement of numerical identity—in the fundamental case, a statement of the kind: the number of <u>F</u>s = the number of <u>G</u>s—is true, if true, in virtue of the very same state of affairs which ensures the truth of the matching statement of one-one correspondence among concepts, and may be known a priori if the latter may be so known.

The way past Benacerraf's dilemma we are commending is, then, in essence extremely straightforward. So long as we can ascertain that lines are parallel, or that concepts are one-one correspondent, there need be no <u>further</u> problem about our knowledge of certain basic kinds of truths about directions and numbers, for all their abstractness. For provided that the concepts of direction and number can be implicitly defined by Fregean abstraction, we can know statements of direction- and numerical-identity to be true just by knowing the truth of the appropriate statements of parallelism among lines and one-one correspondence among concepts. We can do so for the unremarkable reason that the truth-conditions of the former are fixed by stipulation to coincide with those of the latter.[39]

6. Objections and Qualifications

We conclude by responding—more briefly than we should like—to some likely sources of dissatisfaction with this general form of response to the Dilemma.

The Fregean abstractionist may appear vulnerable to attack from two quite different—and indeed doctrinally diametrically opposed—positions. On the one hand, philosophers of a nominalistic persuasion will challenge his entitlement to endorse or stipulate the requisite abstraction principles, when their left hand side syntax is construed—as the Fregean wishes it to be construed—as semantically significant. On the other, philosophers of a more pronounced platonistic tendency will decry his explanation for almost the opposite reason—in their view, the Fregean abstractionist has not rescued anything worth calling platonism at all, but has merely sold out in favour of an Ersatz so dilute in content as to be nugatory. We shall try to explain how we aim to avoid being caught in the ensuing crossfire.

One kind of nominalist will insist that we may accept abstraction principles as explanatory of the truth-conditions of their left hand sides only if we treat the latter as devoid of semantically significant syntax, beyond their involving genuine occurrences of the term or predicate variables α and β.[40] That is, they will admit as legitimate only an austere reading of the equivalences, according to which their instances involve, on their left hand sides, just occurrences of terms or predicates replacing the variables α and β, embedded in the argument places of a single unbreakable predicate whose surface syntactical complexity betokens no genuine semantic structure. So what, on their only acceptable reading, the equivalences accomplish is the introduction of a new, unstructured 2-place predicate—for example:

'thedirectionof…isidenticalwiththedirectionof__' as no more than alternative notation for the equivalence relation that figures on the right hand side—in this case, the equivalence relation expressed by: '… is parallel to __'. The crucial question for a nominalist of this stripe is how he justifies his insistence upon this austere reading. His likely answer will be that since the statement 'lines a and b are parallel' plainly involves no reference to directions, conceived as a species of abstract object distinct from lines themselves, it can be taken as explaining the truth-conditions of 'the direction of line a = the direction of line b' only if the latter is likewise understood as devoid of reference to any such objects. But this seemingly simple answer implicitly relies upon a question-begging assumption. It is certainly true that statements of line-parallism involve no terms purporting reference to directions. And it is equally clearly true that two statements cannot have the same truth-conditions if one of them carries a commitment to the existence of entitities which is absent from the other. But the premiss the nominalist requires—if he is to be justified in inferring that statements of parallelism can be equivalent to statements of direction-identity only if the latter are austerely construed—is that statements of parallelism do not demand the existence of directions. This does not follow from the acknowledged absence, in such statements, of any explicit reference to directions. The statement that a man is an uncle involves no explicit reference to his siblings or their offspring, but it cannot be true unless he has a non-childless brother or sister.

The nominalist may—indeed, must—accept that there can be existential commitment in the absence of explicit reference, but may now object that whilst no one would count as understanding the statement that Edward is an uncle if she could not be brought to agree that its truth requires the existence of someone who is brother or sister to Edward and father or mother of someone else, it is quite otherwise with statements of parallelism. Someone may perfectly well understand the statement that one line is parallel to another without being ready to acknowledge—what the nominalist denies—that its truth requires the existence of directions. The abstractionist reply is that so much is perfectly correct, but insufficient for the nominalist's argument. It is <u>correct</u> precisely because—just as the abstractionist requires—understanding talk of lines and parallelism does not demand possession of the concept of <u>direction</u>. But it is <u>insufficient</u> because the question that matters is, rather, whether one who is fully cognisant with all the relevant concepts—that is, the concepts of line, parallelism and direction—could count as fully understanding a statement of parallelism without being ready to agree that its truth called for the existence of directions. If the concept of direction is implicitly defined, as the abstractionist proposes, by means of the Direction Equivalence, she could not. Insisting at this point that no such explanation can be admitted, because only an austere reading of the Direction Equivalence is permissible, amounts to no more than an unargued—and now explicitly question-begging—refusal to entertain the kind of explanation on offer. This kind of nominalist wants to make out that in laying down an abstraction principle, non-austerely construed, one must be illegitmately assuming the existence of new objects, just as one would be if one were to stipulate the biconditional that the youngest sister of Edward is the youngest sister of Bill if and only if Edward and Bill are brothers. But in his attempt to do so, he simply refuses to take seriously the idea that statements on either side of an instance of an abstraction principle may share their truth-conditions—as is manifestly not the case in the example about sisters and brothers.

The other accusation, that abstractionist platonism falls so far short of the genuine article as to be unworthy of the name 'platonism' at all, is harder to come to grips with, partly because it is hard to find a clear, articulate and non-metaphorical account of what 'genuine' platonism is supposed to involve and partly because any of several distinct things may lie behind it.

One point around which the vague feeling that Fregean platonism is insufficiently 'robust' may crystallise connects directly with the charge—itself emanating from a quite different direction—that the proposal to define number implicitly by means of Hume's Principle amounts to an illegitimate attempt to secure the existence of numbers by stipulation—to define them into existence.[41] The would-be more robust platonist may sympathise with the charge, but see it as highlighting the abstractionist's failure to justify the mind-independence of numbers—instead, they turn out, for the abstractionist, to be mere creatures of our understanding, faint and insubstantial shadows cast by mathematical language as reconstructed on abstractionist lines. The charge itself—and with it, the invidious gloss the robust platonist puts upon it—overlooks at least two vital points. First, what is stipulated, when an abstraction principle is advanced as an implicit definition, is not the existence of certain objects—referents for the terms featured on its left hand side—but the truth of (indefinitely many) biconditionals co-ordinating identity-statements linking such terms with statements involving the relevant equivalence relation over the underlying domain. The truth of any given one of those identity statements, and hence the existence of objects to which its ingredient terms refer, is not stipulated, but follows only given the truth of the co-ordinated statement to the effect that the abstraction's equivalence relation holds among the relevant objects (or concepts, in case of a higher-order abstraction such as Hume's Principle). And the truth of that latter statement will be always a matter of independently constituted fact (about parallelism of certain lines, or one-one correspondence between certain concepts, etc.). What is brought into existence by the stipulation—if anything is—is not objects, but a certain sortal concept. What objects, if any, fall under it is—as we've said—entirely dependent upon the truth of instances of the abstractive biconditional's right hand side. Second, whilst it is a contingent matter of human psychology whether the possibility of forming a sortal concept by means of such a stipulation happens to occur to us, the success of the stipulation is an entirely objective matter—essentially, a matter of its compliance with versions of the constraints—including, centrally, consistency and a certain kind of conservativeness, but probably others besides—which govern definition and concept-formation in general. Once it is seen that the existence of abstract objects is thus essentially dependent upon the objective truth of instances of the right hand side of an abstraction principle, whose acceptability is itself dependent upon its conformity to objective constraints, it is hard to see what can remain of the charge that those objects are somehow deficient in point of mind-independence.

Still, there is a vestigial resistance that may encourage a subscriber to press the charge. In the *Tractatus*, Wittgenstein conceived of the fit between a true statement and that aspect of the world—the fact—which makes it true, as—literally speaking—an isomorphism between the ready-structured fact and the statement concerned. It is traces of this highly realist view that can encourage the vestigial resistance. For suppose we allow the neo-Fregean to offer his best example of a legitimate abstraction principle, meeting whatever other constraints he sees fit to impose. The question then arises: what is the real structure of the type of state of affairs whose obtaining, by his intent, constitutes the shared truth-condition of instances of its left- and right-hand sides? Not both, it seems, can be faithful to that structure, for—so the objection contends—they represent it in conflicting ways. So if it is granted—as is plausible—that the original discourse, supplying vocabulary for the right-hand sides, was perfectly adequate for its purpose, then it appears we must regard the syntax of the language of the left-hand sides as somehow misleading, or at least view the ontological implications of its novel devices of singular reference and quantification as somehow less robust than in discourses where genuine fidelity to predetermined ontological structure is achieved. To be sure, we have introduced a discourse in which claims are now legitimate which sound and behave just like claims of existence and identity. But don't we have to regard them as paling by comparison with the similar-sounding and behaving claims in discourses where a real reflection of the structure of reality is accomplished—including, by hypothesis, the discourse of the right-hand sides?

The objection seems well-taken once its underlying realist metaphysics is allowed (though there is scope for some skirmishing about whether, once in its grip, we can easily dismiss the possibility that it is the left-hand side of the abstraction that is the better ontological guide.) But if the metaphysics is rejected, do we not have to elect instead for the opposing idealist extreme: that the world owes its structure—in particular, its articulation into objects and properties of various kinds—to the forms imposed upon it by our thought? And if that view is intelligible at all, it hardly offers a way of saving the mind-independence of the objects introduced by abstraction, or indeed of any others. So on one horn of the dilemma—realism—the neo-Fregean confronts etiolation of the ontological commitments of the discourse the abstraction serves to introduce; while on the other—idealism—the consequence is mind-dependence.

The solution must be to refuse both horns. The metaphysics of abstractionism must rather be a form of quietism about the realist/idealist alternative that the objector here seeks to impose. We therefore reject the idea, implicit in Tractarian realism, that truth involves a transcendental fit between the structure of reality and the structure of our forms of thought;

but equally we reject the idea that we structure an otherwise amorphous world by the categories in which we think about it. Rather we can merely acquiesce in the conception of the general kinds of things the world contains which informs the way we think and talk and is disclosed in our best efforts to judge the truth. This is all we can justifiably do. For we have no means of independent assessment of the issue of fit; and that unavoidable lack is no reason for idealism. Holding a pack of cards in one's hand really is holding four (token) suits and fifty-two token playing cards. These descriptions are not in competition for true reflection of the structure of the world. But to recognise that is hardly to fall into idealism. So with the parallelism of a pair of lines and the identity of their directions.

We are under no illusion, of course, that these brief remarks constitute a fully adequate treatment of the issues here. Our point is merely that what may once again seem a quite intuitive objection to the neo-Fregean approach is actually rooted in anything but intuitive metaphysical assumptions.

We have been exclusively concerned with one central aspect of Benacerraf's dilemma, as it confronts the platonist—the thought that the very abstractness of the objects of which, as the platonist construes them, mathematical statements speak must, in and of itself, render mathematical knowledge impossible. We have tried to explain how, if talk and thought of abstract objects is grounded in abstraction principles such as Hume's Principle, taken as implicitly definitional of relevant sortal concepts such as cardinal number, this apparent epistemological impasse may be circumvented.[42] But there are many other aspects of the problem of mathematical knowledge, even in the most elementary case of the arithmetic of natural numbers. Many arithmetic truths, including most mathematically interesting ones, depend essentially on the <u>infinity</u> of the sequence of natural numbers. Part of the problem of mathematical knowledge is therefore to explain how we can know that there is such a sequence and know truths about it. This problem does not have anything especially to do with the <u>abstractness</u> of the natural numbers—save, perhaps, in the minimal sense that it is probably only in the case of abstract objects of some kind that we can expect to know, or at least be able to prove, that there exists an infinity of them—so it is not a problem that distinctively afflicts platonism, but a problem for everyone. But it demands a solution, and the Fregean platonist is—in our view— nicely placed to provide one. For, as is well-known, the adjunction of Hume's Principle to a suitable underlying second-order logic enables us to prove the five Dedekind-Peano axioms for arithmetic. The proof[43] depends crucially on a feature of Hume's Principle which sets it apart from simpler abstractions such as the Direction Equivalence, namely its impredicativity. The Fregean must therefore defend himself against the charge that this impredicativity amounts to vicious circularity or is somehow otherwise objectionable.[44]

And of course whilst there is less than universal consensus—even among the classically minded—about the full extent of our mathematical knowledge, no one supposes it stops short at elementary arithmetic! A viable epistemology for mathematics ought to encompass, at a minimum, the foundations of the theory of real numbers, the various branches of analysis, and, arguably, at least a significant portion of set theory, even if not the whole of it. From the Fregean perspective, the question is whether we can formulate otherwise acceptable abstraction principles of sufficient power for the recovery of these theories.[45] If that can be done, extending the abstractionist programme, and its answer to Benacerraf's Dilemma, beyond arithmetic will, so far as we can see, raise no new epistemological issues.[46]

Bob Hale
Glasgow University

Crispin Wright
University of St. Andrews and New York University

[N]OTES

[1] Benacerraf (1973). All page references to this article are to the reprint in Benacerraf and Putnam (1983).

[2] –involving (numerically definite) quantification over objects, one- and two-place predicates of first-level (i.e. true or false of objects) and singular terms for objects, all put together in the same way.

[3] The general problem of providing, for a given area of thought and talk, an account of the truth-conditions of its distinctive claims which simultaneously satisfies both requirements is what Christopher Peacocke calls the Integration Challenge. See Peacocke (1999), ch.1. As Peacocke emphasises, the Integration Challenge is, in effect, a generalisation of Benacerraf's problem.

[4] For which see Field (1980) and (1989), essays 1,2,5 and 7.

[5] Hellman (1989)

[6] The quite extensive critical literature on Field's unorthodox variety of nominalism includes the important review of Field (1980) by Malament (1982), Shapiro (1983), Urquhart (1990), Maddy (1990), ch.5, and Chihara (1990), chs.8 and12. Critical reactions of our own are given in Hale (1987) and (1990), and Hale & Wright (1993) and (1994).

Hellman's approach is one of the major strategies explored in Burgess and Rosen (1997). Critical appraisals may be found in Resnik (1992), Shapiro (1993) and Hale (1996).

On Quine's indispensability argument, there is a vast mass of literature. Field's own position is a reaction to the thought that if platonism could be justified at all, Quine's way would be the only way. Parsons (1983) ch.7 and Maddy (1997) Pt.II, ch.2 & Pt.III, ch.3 are especially helpful.

[7] The association is largely based on the following well-known passage:

'But, despite their remoteness from sense-experience, we do have something like a perception of the objects of set theory, as is seen from the fact that the axioms force themselves upon us as being true. I don't see any reason why we should have less confidence in this kind of perception, i.e. in mathematical intuition, than in sense-perception.' (Gödel (1947), pp.483-4)

It's open to question whether this passage really sustains its widely accepted non-naturalist reading. SeeW.W. Tait (1986), n.3, p. 147 and following.

[8] cf Parsons (1980), sections I-III

[9] This is, in essence, Dummett's characterisation, given in Dummett (1973), ch.14, p.503. Parsons prefers to draw the distinction in different terms, but agrees that there is extensional coincidence–(at Parsons (1999), I. p. 49)

[10] cf. Parsons (1980), p103. One must, of course, take the class of stroke-strings (i.e. stroke-string types) to be closed under the operation of adding a stroke.

[11] Parsons (1980), p.103

[12] Parsons (1980), p.103

[13] Cf Parsons (1993) p.240

[14] Parsons (1980), p.104

[15] Parsons (1980), p.104

[16] Parsons does not himself explicitly formulate them in this way. We are following James Page's (1993) exposition. But Parsons evidently has no quarrel with Page's attribution to him of the view that (PA1′-4′) are intuitively known–at least, he takes no exception to it in his reply to Page in Parsons (1993).

[17] Cf Parsons (1980), p.105

[18] Parsons (1980), p.106

[19] Ibid. As Parsons also emphasises, it is essential, if the thought experiments are to confer the requisite generality, that they be carried out on the basis of specific concepts, such as that of a string of strokes.

[20] We are not attributing this proposal to Parsons. While he certainly evinces sympathy with a version of structuralism, he rejects both that form of it—what Dummett terms it, 'mystical' structuralism and Shapiro 'ante rem' structuralism—which takes mathematical theories to be about abstract structures and also eliminative structuralism, even in its modal versions. See Parsons (1990) and (1999), chs. 2, 3. We are not clear what Parsons' preferred form of structuralism comes to, or how it connects with the claims he wishes to make on behalf of mathematical intuition.

[21] In which $\underline{A}'$ is a suitably adjusted version of our original arithmetic statement $\underline{A}$, incorporating a free variable place for variables ranging over, inter alia, ω-sequences.

[22] —contrary to what Hellman seems to suppose. For detailed argument in support of this claim, see Hale (1996)

[23] At least, he denies that one can have intuitions of numbers, and we cannot see how his reasons for denying this won't apply to pure abstract objects in general.

[24] Shapiro's views about the nature of mathematical ontology have, of course, an important ally in Michael Resnik—see especially the latter's (1997). For our purposes, however, the crucial differences between these two thinkers lie in their respective epistemological views.

Resnik's epistemology is fundamentally holistic and Quinean and the ultimate ground for our mathematical knowledge, even construed as concerning pure structures, is no different to that for science in general. He is therefore no conservative in our sense. Shapiro's approach, by contrast, is hospitable to the apriority of classical mathematics.

[25] Shapiro (1997), p.115)

[26] Shapiro (1997), p.119:

> "... We first contemplate the finite structures as *objects* in their own right. Then we form a *system* that consists of the collection of these finite structures with an appropriate order. Finally, we discuss the *structure* of this system. Notice that this strategy depends on construing the various finite structures, and not just their members, as *objects* that can be organised into systems. It is *structures* that exhibit the requisite pattern."

The reader will note the very different conception Shapiro expresses of the roots of our recognition of simply infinite structures from that evinced in Parsons. For Shapiro, recognising that there are infinitely many stroke-strings is <u>holistic</u>, based on the appreciation of a second order pattern among suitably arranged patterns exemplified by single stroke-string tokens. For Parsons, by contrast, the recognition proceeded on the basis of a single intuition of an arbitrary stroke string type, given either in perception and imagination, and an apprehension—the point we found problematical above—that it could be extended by one more stroke.

[27] As Shapiro is, of course, well aware, applied to the equivalence relations on the objects in a given domain, this ... will not provide for more abstracts than there are such objects in any case where the domain is infinite. He therefore envisages its application to the equivalence relations on the properties over a given domain, making the standard assumption that these will number 2^n for any n-fold domain.

[28] Benacerraf himself hints (p.412) that the problem for the platonist is as much a problem about reference as about knowledge. Field is more explicit on the matter, suggesting that the problem of reference is the more fundamental–see, for example, his remarks on the matter in his (1982), p.68.

[29] Frege (1884)

[30] Obviously not just any statement incorporating numerical terms will do. The existence of a number would not follow from the truth of '$\neg\exists x$ x = the number of Jupiter's moons'. There is more to say but it will serve the immediate purpose to restrict attention to atomic statements.

[31] The 'acquaintance first' conception goes naturally with and supports the idea, explicitly at odds with the context principle, that word-meaning is somehow prior to sentence meaning and the corresponding primacy accorded by traditional empiricist philosophy to ostensive definition in the acquisition of understanding of terms

and predicates. The relevant cluster of misconceptions about ostensive definition take centre stage at Philosophical Investigations §§28-36. For further discussion, see Wright (1983), ch.1, sections vii and viii

[32] Following the suggestion in Boolos (1990), p.268

[33] Frege (1983), §64

[34] Frege (1892), p.47.

[35] For further discussion, see Dummett (1991), ch.14, and especially p.173ff, and (1981), ch.15; Linsky (1992); Hale (1997) and (2001); Potter & Smiley (2001)

[36] It is of crucial importance here that sameness of truth-conditions may be understood as not requiring identity of sense under the strongly compositional interpretation of sense Frege assumes, according to which the senses of the parts of a complex expression must be parts to the sense of the whole (cf. Frege (1893), §32).

[37] Cf. Frege (1884), §22 and §46.

[38] See Frege (1884), §64, final paragraph, and §§104-5

[39] Of course, accounting for knowledge of statements of identity is but a first step. The abstractionist has next to show how a an epistemological foundation for an entire mathematical theory—arithmetic, analysis, or even set-theory—may be built upon this first step.

[40] Some traditional nominalists will, in addition, jib at the quantification through property variables involved in second order abstractions such as Hume's Principle. The neo-Fregean abstractionist's right to take second-order logic as in good standing is, of course, an important and controversial issue, but it is a further issue, separable from the one that concerns us here, and we must set it aside here.

[41] For one recent expression of this charge, see Potter & Smiley (2001), and for a response, Hale (2001).

[42] As anyone at all familiar with the territory will know, there are many difficulties for and objections to the Fregean approach which we have not been able to mention, much less address. For a survey, see the Introduction and Postscript to Hale & Wright (2001). A particularly useful overview of the issues is provided by Fraser MacBride (forthcoming)

[43] Prefigured by Frege in (1884), §§82,83 and reconstructed by Wright in Wright (1983), §xix. Other detailed accounts of the proof are given by George Boolos in Boolos (1987), in an appendix to Boolos (1990), and in Boolos & Heck (1998)

[44] Dummett (1991), ch.18; Hale (1994); Wright (1998a); Dummett (1998);Wright (1998b)

[45] There is increasing evidence for an affirmative answer as far as Real Analysis is concerned. See Hale (2000). Alternative, and in some respects simpler, ways of introducing the reals by abstraction are developed by Stewart Shapiro and Crispin Wright, and by Jeffrey Ketland, in forthcoming papers.

[46] Some of the material in this paper was originally presented by Bob Hale at the Conference on the Epistemology of Basic Belief, held in the Free University of Amsterdam, June 20th —22nd 2001. We thank Fraser MacBride for helpful suggestions and discussion and Charles Parsons for help in clarifying certain aspects of his position and permission to cite his (1999). This paper was composed during Crispin Wright's tenure of a Leverhulme Research Professorship; he gratefully acknowledges the support of the Leverhulme Trust.

REFERENCES

Benacerraf, Paul (1973) 'Mathematical Truth' Journal of Philosophy 70 (1973), pp.661-80.

Benacerraf, Paul & Putnam, Hilary (Eds.) (1983) Philosophy of Mathematics: Selected Readings 2nd edn. Cambridge: Cambridge University Press.

Boolos, George (1987) 'The Consistency of Frege's Foundations of Arithmetic' originally published in J.Thomson, ed. On Being and Saying: Essays in Honor of Richard Cartwright Cambridge, Mass: MIT Press 1987, pp.3-20, and reprinted in Boolos (1998), pp.183-201 and in Demopoulos (1995), pp.211-233)

(1990) 'The Standard of Equality of Numbers' in George Boolos, ed. Meaning and Method: Essays in Honor of Hilary Putnam Cambridge: Cambridge University Press (1990), and reprinted in Boolos (1998), pp.202-19 and in Demopoulos (1995), pp.234-54

(1998) Logic, Logic and Logic Cambridge MA: Harvard University Press

Boolos, George & Heck, Richard Jnr. (1998) 'Die Grundlagen der Arithmetik §§82-3' in Schirn (1998), pp.407-28, and reprinted in Boolos (1998), pp.315-38

Burgess, John and Rosen, Gideon (1997) A Subject With No Object Oxford: Clarendon Press

Chihara, Charles (1990) Constructibility and Mathematical Existence Oxford: Clarendon Press

Demopoulos, W (Ed.) (1995) Frege's Philosophy of Mathematics, Cambridge,MA: Harvard University Press

Dummett, Michael (1981) The Interpretation of Frege's Philosophy London: Duckworth

(1991) Frege: Philosophy of Mathematics London: Duckworth

(1998) 'Neo-Fregeans: in Bad Company?' in Schirn (1998), pp.369-87

Field, Hartry (1980) Science without Numbers Oxford: Basil Blackwell

(1982) 'Realism and anti-realism about mathematics' Philosophical Topics 13, pp.45-69, reprinted in Field (1989). Page references are to this reprint.

(1989) Realism, Mathematics and Modality Oxford: Basil Blackwell

Frege, Gottlob (1884) Die Grundlagen der Arithmetik Breslau: Wilhelm Koebner, translated into English by J.L.Austin as The Foundations of Arithmetic Oxford: Blackwell 1959. Page references are to this translation.

(1892) 'Über Begriff und Gegenstand', originally published in the Vierteljahrsschrift für wissenschaftliche Philosophie 16, pp.192-205, and translated as 'On Concept and Object' in Peter Geach & Max Black's Translations from the Philosophical Writings of Gottlob Frege Oxford: Blackwell 1952, pp.42-55. Page references to this translation.

(1893) Die Grundgesetze der Arithmetik vol.1 Jena: H.Pohle, part translated into English by Montgomery Furth in The Basic Laws of Arithmetic Berkeley: University of California Press 1964

Gödel, Kurt (1947) 'What is Cantor's Continuum Problem?' originally published in American Mathematical Monthly 54, pp.515-25; revised version in Paul Benacerraf & Hilary Putnam, eds. Philosophy of Mathematics: Selected Readings 2nd edn. Cambridge: Cambridge University Press 1983, pp.470-85. Page references are to this reprint.

Hale, Bob (1987) Abstract Objects Oxford: Basil Blackwell

(1990) 'Nominalism' in A.D.Irvine(1990) p.121-44

(1994) 'Dummett's critique of Wright's attempt to resuscitate Frege' Philosophia Mathematica (3) Vol 2, pp.122-47, reprinted in Hale & Wright (2001), pp.189-213

(1996) 'Structuralism's unpaid epistemological debts' Philosophia Mathematica (3) Vol.4, pp.124-147

(1997) 'Grundlagen §64' Proceedings of the Aristotelian Society 97, pp.243-61, reprinted with a new appendix in Hale & Wright (2001), pp.91-116.

(2000) 'Reals by Abstraction' Philosophia Mathematica (3), 8, pp.100-23

(2001) 'Reply to Potter and Smiley' Proceedings of the Aristotelian Society 101, pp.339-358.

Hale, Bob & Wright, Crispin

(1993) 'Nominalism and the Contingency of Abstract Objects' Journal of Philosophy Vol.89, No.3, pp.111-135

(1994) 'A reductio ad surdum? Field on the contingency of mathematical objects' Mind Vol 103, No 410, pp.169-184

(2001) The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics Oxford: Clarendon Press

Hart, W.D. (Ed.) (1996) The Philosophy of Mathematics Oxford: Oxford University Press

Hellman, Geoffrey (1989) Mathematics without Numbers Oxford: Clarendon Press

Irvine, Andrew (Ed.) (1990) Physicalism in Mathematics Dordrecht: Kluwer Academic Publishers,

Maddy, Penelope (1990) Realism in Mathematics Oxford: Clarendon Press

(1997) Naturalism in Mathematics Oxford: Clarendon Press

Malament, David (1982) 'Review of Field (1980)' Journal of Philosophy 79, pp.523-34

Linsky, Bernard (1992) 'A Note on the'Carving Up Content' Principle in Frege's Theory of Sense' Notre Dame Journal of Formal Logic 33, pp. 126-135

Page, James (1993) 'Parsons on Mathematical Intuition' Mind Vol.102, No.406, pp.223-32

Parsons, Charles (1980) ‘Mathematical Intuition’ Proceedings of the Aristotelian Society 80, pp.145-68, reprinted in W.D.Hart, ed. The Philosophy of Mathematics Oxford: Clarendon Press (1996). Page references are to this reprint.

(1983) Mathematics in Philosophy: Selected Essays Ithaca, New York: Cornell University Press

(1986) ‘Intuition in Constructive Mathematics’ in J.Butterfield, ed. Language, Mind and Logic Cambridge: Cambridge University Press, 1986, pp.211-229

(1993) ‘On Some Difficulties Concerning Intuition and Intuitive Knowledge’ Mind Vol.102, No.406, pp.233-46

(1998) ‘Finitism and Intuitive Knowledge’ in Schirn (1998), pp.249-70

(1999) Mathematical Thought and its Objects [unpublished typescript]

Peacocke, Christopher (1999) Being known Oxford : Oxford University Press

Potter, Michael & Smiley, Timothy (2001) ‘Abstraction by Recarving’ Proceedings of the Aristotelian Society 101, pp.327-38

Resnik, Michael (1992) ‘AStructuralist's Involvement with Modality’ Mind 101, pp. 107-22

(1997) Mathematics as a Science of Patterns Oxford: Clarendon Press

Schirn, Matthias (ed.) (1998) Philosophy of Mathematics Today Oxford: Clarendon Press

Shapiro, Stewart (1983) ‘Conservativeness and incompleteness’ Journal of Philosophy 80, pp. 521-31

(1993) ‘Modality and Ontology’ Mind 102, pp. 455-81

(1997) Philosophy of Mathematics: Structure and Ontology New York & Oxford: Oxford University Press

Tait, W. (1986) 'Truth and Proof: the Platonism of Mathematics' Synthese 69, pp. 341-70; reprinted in Hart (1996) pp. 142-67. Page references are to the latter volume.

Urquhart, Alasdair (1990) 'The Logic of Physical Theory' in Irvine (1990) pp. 145-54

Wright, Crispin (1983) Frege's Conception of Numbers as Objects Aberdeen: Aberdeen University Press

(1998a) 'On the Harmless Impredicativity of $N^{=}$ (Hume's Principle)' in Schirn (1998), pp.339-68, reprinted in Hale & Wright (2001), pp.229-55

(1998b) "Response to Dummett" in Schirn (1998), pp.389-405, reprinted in Hale & Wright (2001), pp.256-71