The revival of the philosophy of mathematics in the 60s following its post-1931 slump left us with two conflicting positions on arithmetic's ontological relationship to set theory. W.V. Quine's view, put forward in *Word and Object* (1960), was that numbers are sets. The opposing view was advanced in another milestone of twentieth-century philosophy of mathematics, Paul Benacerraf's "What Numbers Could Not Be" (1965): one of the things numbers could not be, it explained, was sets; the other thing numbers could not be, even more dramatically, was objects. Curiously, although Benacerraf's article appeared in the heyday of Quine's influence, it declined to engage the Quinean position squarely, even seemed to think it was not its business to do so. Despite that, in my experience, many philosophers believe that Benacerraf's article put paid to the reductionist view that numbers are sets (though perhaps not the view that numbers are objects). My chapter will attempt to overturn this orthodoxy.

1 Reductionism

Benacerraf observes that there are many potential reductions of arithmetic to set theory. The most familiar are the "Ernie" (von Neumann) interpretation, which takes 0 as the empty set and the successor of x as $x \cup \{x\}$, and the "Johnny" (Zermelo) interpretation, which also takes 0 as the empty set but the successor of x as $\{x\}$. With New World irreverence, the set theorists' first names have been switched round and Americanised. No two such accounts can be correct (since e.g. 2 cannot be equal to both $\{\emptyset, \{\emptyset\}\}$ and $\{\{\emptyset\}\}$), and since there is no principled way to choose between them, "any feature of an account that identifies a number with a set is a superfluous feature of the account (i.e. not one that is grounded in our concept of number)" (Benacerraf 1998, p. 52). Hence, numbers are not sets.

Our aim here is to defend reductionism from Benacerraf's and other arguments. Our main defence will be that speakers might not have transparent knowledge of the referents of their number terms, and that any concern with getting actual meanings right must be subordinate to the overall theoretical virtues of a proposed account of arithmetic. We first outline some of reductionism's varieties: two objectual versions, a structural version and a partial-denotation version. These reductionisms all share a commitment to an ontology that includes sets but not *sui generis* numbers (i.e. numbers as an independent kind of thing). They differ only on the semantics for arithmetic.

1.1 Objectual reductionisms

 Objectual reductionism is the doctrine's classic form, and the one Benacerraf had in mind when presenting his argument. Objectual reductions associate with each number a set and interpret the arithmetical claim "$\varphi_A(N_1, N_2,...)$" as the corresponding set-theoretic claim "$\varphi_S(S_1, S_2, ...)$". For example, "$0 < N_1$" and "$N_1 + 1 = N_1$" are interpreted as "$\emptyset \in S_1$" and "$S_1 \cup \{S_1\} = S_1$" on the von Neumann interpretation. A successful reduction shows that

numbers are sets, and that claims about arithmetic are nothing other than claims about (some) sets. Objectual reductions come in two flavours: context-independent and context-dependent. The former identifies numbers with some sets once and for all, independently of context. The latter identifies numbers with different sets in different contexts, for example, von Neumann finite ordinals in the context of one proof, Zermelo ordinals in the context of another.[1] On some people's terminology, only objectual reductionisms truly deserve the label "reductionism"—see below.[2]

1.2 Structural reductionism

Rather than picking a particular reduction (in all contexts or context-dependently), structuralists quantify over all set-theoretic omega-sequences. A semi-formal interpretation of the (false) arithmetic statement "$0 = 1$", for example, would be: "for any set X, for any set-theoretic function S_X on X and member 0_X of X, if the Peano axioms hold of the tuple (X, S_X, 0_X), then $0_X = S_X(0_X)$". This type of reductionism is sometimes known as set-theoretic structuralism.[3,4]

1.3 Partial-denotation reductionism

A partial-denotation reduction attempts to combine objectual semantics with the idea that arithmetical terms can stand for the entities they would partially stand for in the von Neumann objectual reduction and the Zermelo one and many others. A set-theoretic interpretation partially accords with arithmetical language if each term partially denotes the entity the interpretation assigns to it; and there are many such interpretations. Thus, the ω-

---

[1] The Quine of *Word and Object* is best interpreted as a context-dependent objectual reductionist (1960, p. 263).

[2] If concepts are distinguished from their extensions, then objectual reductionists need not claim that our arithmetical *concepts*, either pre- or post-reduction, are set-theoretic. For example, context-independent objectual reductionists can maintain that the concept of zero is different from the concept of the empty set even if they have the same extension.

[3] Do not confuse this with the view that set theory itself is to be interpreted structurally.

[4] Did Benacerraf himself adopt structural reductionism in "Way out", the punnily titled final section of his "What Numbers Could Not Be" (1965)? It is hard to tell for several reasons. (i) That section is obscure and its argument can be interpreted in many ways. See Wetzel (1989). (ii) Benacerraf does not seem to privilege set theory over number theory, so parity of reasoning would then seem to commit him to a structural interpretation of both, which is definitely not the structuralism proposed above. (iii) As we will see, structural reductions have to choose between taking ordered tuple relations as primitive or arbitrarily reducing them to sets, an analogue of his own multiplicity objection against objectual reductionism. Benacerraf does not so much as hint at this problem, which suggests that this is not what he had in mind. (iv) Benacerraf (1998, p. 50) later signalled that "Way out" was primarily intended as meaning "crazy" rather than "solution". For all these reasons, it is hard to say with confidence that structural reductionism is Benacerraf's preferred way out.

sequence of numerals partially denotes several ω-sequences of sets.[5] Field (1974) is the original source for partial-denotation semantics, for mathematics and more generally. He argues there that "gavagai" might partially denote rabbit and undetached rabbit part; pre-relativistic tokenings of "mass" might partially denote relativistic mass as well as rest mass, and so on.

What is the motivation for reductionism? It tends to be economy, of various stripes. First, ontological economy: the reductionist believes in fewer things and, depending on her view, fewer types or fewer higher-level types of things than the non-reductionist, since she believes in sets rather than sets and *sui generis* numbers.[6] Second, ideological economy: reductionism reduces the number of primitive predicates our total theory employs. Third, axiomatic economy: the list of basic principles is reduced or simplified.[7] Reductionists also sometimes cite other motivations, for example that their view avoids or answers vexing inter-categorial identity questions such as whether the natural number 2 is identical to the rational number 2 or the real number 2 or the complex number $2 + 0i$.[8] Our interest is in whether reductionism can rebut Benacerraf's arguments, so we shall not evaluate these motivations but take them as given. Reductionism as here understood is inspired by Quine but is free from Quinean idiosyncracies. Quine developed his set theory NF as a rival to iterative set theories such as Zermelo–Fraenkel set theory with Choice (ZFC) and von Neumann–Bernays–Gödel class theory (NBG). Indeed, he believed that the only intrinsically plausible conception of set is the naïve one and that, following the discovery of the set paradoxes associated with that conception, the choice of set theory is purely pragmatic. Reductionism is not committed to this view nor to any particular reducing set theory, so long as it is stronger than arithmetic.[9] Reductionists also need not follow Quine in his scientific naturalism, which takes science to

---

[5] For brevity, I speak here and elsewhere of numerals or numerical terms denoting sets. Of course, numerical predicates, function symbols and quantifiers are also to be understood as their set-theoretic counterparts.

[6] For reductionists who take the sets corresponding to the domain of arithmetical discourse to be a distinct kind, a qualitative economy of higher-level kinds is achieved via the kind inclusion of natural numbers to sets. For reductionists who do not, the kind *number* has been eliminated. For Quine's related distinction between what he calls explication and elimination, see (1960, §§53–5, esp. p. 265).

[7] Reduction and/or simplification may result even if the size of the axiom set is unchanged, for example, (i) by accepting (countably many) instances of one schema rather than (countably many) instances of two schemas, (ii) by accepting only A rather than A ∧ B.

[8] "A reduction clears out Plato's slum, ridding it of its 'primeval disorder'" Quine (1960, p. 267). Briefly, for the objectual reductionist, all these entities are sets and the identity claims are settled by the various chosen semantics for the theory of natural numbers, rationals, reals and complex numbers. For the structuralist, a claim such as "$2_N = 2_Q$" would be understood along the lines of "for any system of sets S_N doing the job of the natural numbers and any system of sets S_Q doing the job of the rationals, 2-in-the-S_N-system = 2-in-the-S_Q-system", which is false. Likewise, this statement's negation "¬$2_N = 2_Q$" would also be false under the structuralist interpretation. As the structural reductionist sees it, this is not a worry since mathematics is not firmly committed to bivalence for intra-theoretic statements of this kind. (See fn. 15 for a related point.) The case of the partial-denotation reductionist is similar to that of the structural reductionist, *mutatis mutandis*.

[9] Or, perhaps so long as it is at least as strong as arithmetic. In that case, the reduction of arithmetic to set theory might not be preferable to that of set theory to arithmetic.

be authoritative and accordingly sees mathematics as true only to the extent that it is indispensably applied. Nor need reductionism be committed to ontological relativity, indeterminacy of translation, semantic behaviourism and holism, the thesis that first-order logic is the only true logic, and Quine's sometimes deflationary conception of ontology. The only Quinean doctrine, now widely shared, reductionists must unarguably accept is that their stated motivations are epistemic and can guide the choice of total theory.[10]

2 Reductionism's response

How do our four reductionisms handle Benacerraf's argument? The argument is aimed at reductions that aspire to reveal what number-talk meant all along, which we might call *meaning analyses*. There is nothing in our current concept of number to favour the von Neumann over the Zermelo reduction, and since both cannot be true, neither is. An explication, in contrast, is immune to this problem. Given that set theory can do the job of arithmetic, we can dispense with a commitment to *sui generis* numbers. That there is nothing in our current concept of number to favour one reduction over another does not detract from an explication's desirability, since explications are not constrained to respect our current concept. What remains is the minor issue of arbitrariness in the semantics. Let us see how this plays out for each reductionism.

2.1 Context-independent objectual reductionism

The decision to choose any reduction over any other is arbitrary. But that is a theoretical deficit so minor, or perhaps not a deficit at all, that it is overridden by the greater benefit of avoiding commitment to sui generis numbers, reducing the stock of one's ideological primitives, and so on. Compare the story of Buridan's ass (or rather Aristotle's man—see *De Caelo* 295b). This unfortunate creature died of inanition as a result of not breaking the tie between a bucket of water to its left and a stack of hay an equal distance to its right, at a time when it was equally hungry and thirsty. The two choices were perfectly symmetric, and there was no reason to prefer one to the other. But there was reason to choose one of them. Rationality counsels Buridan's ass to break the tie arbitrarily: in this situation, it is more rational to make an arbitrary choice than no choice. The story of Buridan's ass is, I hope, imaginary, but that does not matter, and if you like your examples a touch more realistic, observe that the choice faced by Buridan's ass is approximated on a daily basis by shoppers in a supermarket choosing between multiple versions of the same good; though in their case, quite reasonably, the paralysis of choice tends to be shorter-lived and less fatal. Similarly, as the reductionist sees it, it is more rational to make an arbitrary selection between two equally good set-theoretic objectual semantics than to refuse to reduce arithmetic to set theory. The benefits of adopting some reduction or other outweigh the arbitrariness of choosing any particular one. Multiple reducibility is therefore not a significant obstacle for this kind of reductionist, just as the multiplicity of choice should not have been for Buridan's ass.

---

[10] Benacerraf seems to think that (confirmatory) holism is an essential part of reductionism (1998, pp. 54–56), but this is not clear. If holism is required, it need not be the strong holism promulgated by Quine that takes our total theory as the unit of confirmation.

For a semantic analogy, suppose a mother of identical twin girls is delivered of them at exactly the same moment.[11] She had long ago settled on "Olympia" for her first-born girl's name and "Theodora" for the second-born; but as they were born simultaneously, she is in a pickle. What to do? An arbitrary choice is called for: say, name the baby on the left "Olympia" and the one on the right "Theodora". Given that each must have a name—what a pity to go through life nameless—it is rational for her to make an arbitrary decision. This decision is purely semantic and no "ontological" choice is involved; the mother is simply choosing one name assignment over another for some given entities, persons in this case. As above, rationality recommends arbitrarily breaking the tie with a terminological decision. Similarly, objectual reductionists should arbitrarily break the tie and fasten on to a particular assignment of numerals to sets.

One might object that the analogy between the number case and the twins case is imperfect, in particular that number terms have an established antecedent usage in a body of claims taken to be truths whereas in our example "Olympia" and "Theodora" do not. The analogy is of course imperfect—it does not hold in several respects. (Yet, notice that the twins story could be told in such a way that there is an established body of truths containing "Olympia" and "Theodora" prior to their referents' birth.) In Section 3, we shall consider a version of this objection based on the fact that number terms have a well-established antecedent use.

It goes without saying that reductionists (objectual or otherwise) need not advocate any surface change to normal mathematical practice. Their claim is not that mathematicians must have the set-theoretic reduction explicitly in mind when doing mathematics, but that they should be committed to it. It is a philosophical claim about mathematics, not a prescription about quotidian practice—though that is compatible with its also being "hermeneutic" (meaning-preserving), as we shall see.[12]

2.2 Context-dependent objectual reductionism

Here, the choice of which reduction to use is seen as turning on pragmatic aspects of the context. For example, if finite arithmetic is considered as a fragment of transfinite arithmetic, von Neumann semantics is preferable to Zermelo's, since the former but not the latter extends neatly into the transfinite. For most contexts, some arbitrariness still remains in picking a particular semantics (e.g. extendability into the transfinite does not uniquely privilege the von Neumann semantics). As with the context-independent version, proponents of this view maintain that this arbitrariness is non-existent or negligible compared to the gain in reducing arithmetic to set theory—a gain in ontological, ideological and axiomatic economy, the avoidance of inter-categorial identity problems, and so on.

2.3 Structural reductionism

On the surface, it seems that a structural reduction, unlike objectual ones, avoids arbitrariness. We need no longer arbitrarily choose set-theoretic representatives for the numbers; we can instead quantify over all omega sequences. However, any structural reduction must reduce the ordered pair relation (more generally, ordered tuple relations) to set

---

[11] Ignore the fact that this is biologically unrealistic.

[12] Note that on the objectual reductionist view, a vexing question of inter-categorial identity turn into the question of whether different branches of mathematics have overlapping interpretations, something to be settled by pragmatic considerations.

theory in some arbitrary fashion. For consider the successor relation *S*'s set-theoretic representation: the reducing set must contain all and only ordered pairs whose first element is a member of *S*'s domain and whose second element is the first's image under *S*. But there is no unique way of cashing out $\langle a, b \rangle$ in purely set-theoretic terms. One can stipulate that $\langle a, b \rangle$ is $\{\{a\},\{a, b\}\}$, or $\{\{\{a\}\},\{a, b\}\}$, or …. This issue recurs with the question of how to explicate higher tuples in terms of pairs. For example, is the triple $\langle a, b, c \rangle$ to be equated with $\langle a, \langle b, c \rangle \rangle$ or $\langle \langle a, b \rangle, c \rangle$?[13,14]

The structural reductionist, therefore, also accepts some arbitrariness, assuming she prefers to reduce mathematics to set theory rather than set theory plus a primitive theory of *n*-tuples (and, given her reductionist tendencies, she presumably does). She relocates rather than avoids the arbitrariness. However, as before, she sees any loss incurred by terminological arbitrariness as outweighed by the reduction's substantive gains.

2.4 Partial-denotation reductionism

A partial-denotation reduction is in the same boat as a structural reduction. Though it avoids choosing one particular ω-sequence of sets as the denotation of the numerals, it avails itself of ordered n-tuples, indeed of infinite-tuples, since it claims that $\langle$ "0", "1", "2", … $\rangle$ partially denotes the von Neumann sequence $\langle \emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots \rangle$ and partially denotes the Zermelo sequence $\langle \emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots \rangle$ and many others. It therefore has to settle on some arbitrary reduction of the various ordered tuple relations to set theory. Once more, the resulting benefits are deemed to outweigh the arbitrariness of choice.

3 Objections

The question of which of the four reductionisms is to be preferred is not one we shall investigate. It is secondary to the ontological question of whether to be a reductionist in the first place. In the assessment of reductionism as a generic position, the debate between the various semantic proposals is internecine. Of course, if it turns out that each species is

---

[13] Cp. Kitcher (1978). You might think that in the case of the ordered pair relation, the reduction of $\langle a, b \rangle$ to $\{\{a\},\{a, b\}\}$ is now so standard that it represents a correct meaning analysis. However, consider this: (i) Even if $\langle a, b \rangle$ is sometimes defined as $\{\{a\},\{a, b\}\}$, this definition is usually acknowledged to be one several possible ones. It is rare that a teacher or text claims it as the meaning of "ordered pair". (ii) The definition is rarely given outside set theory. Mathematicians learn about ordered pairs and ordered n-tuples long before they learn about this reduction, if they ever do. (iii) The arbitrariness of reducing *n*-tuples for $n > 2$ remains; here, there is no conventional reduction of choice.

[14] What if one tries to quantify over all ordered pair reductions? The natural way to do so would be to translate "…$\langle a, b \rangle$…" as "$\forall f(\Phi(f) \to \ldots f(a, b)\ldots)$" where "$\Phi(f)$" abbreviates the claim that *f* is an ordered pairing function. However, in first-order set theory to specify what it is for something to be a function seems to require specifying what it is to be an ordered pair, e.g. one of the conjuncts in the specification of "*f*: $D \to R$" is

$$\forall x[x \in f \leftrightarrow \exists d \in D \exists r \in R(x = \langle d, r \rangle)].$$

(This problem disappears in second-order set-theory, in which functional variables and quantification over them are given.) In any event, the translation gets the truth-conditions of ordered pair statements wrong, for example, the translation of "$\langle a, b \rangle = x$" is false for any *x*.

unattractive, so is the genus. However, that is not the case. As the reductionist sees it, the deficiencies associated with the most attractive reductionism are relatively minor and are outweighed by reductionism's general advantages. Arguably, only the two objectual reductions count as reductions proper, because only they respect arithmetic's surface semantics. Even so, philosophies who reject *sui generis* numbers and take arithmetical truths to be about sets clearly form some sort of philosophical natural kind. We shall hereon concentrate on reductionism's first version, its classic variant, and defend it against objections.

3.1 Semantic objections

Hartry Field sketches two arguments against objectual reductions in the article in which he sets out partial-denotation semantics. The first does not apply to our case of interest,[15] so we focus on his second argument, which does explicitly address it:

> The reductivist might try to escape this difficulty [Benacerraf's multiplicity problem] by saying that it is not important to his purposes to hold that number-theoretic terms referred to these correlated objects all along; it is sufficient (he might say) that we be able to replace number-theoretic talk by talk of the correlated objects. (Quine has suggested this line of escape in 1960, sections 53 and 54.) But this will not do, for it suggests that earlier number theorists such as Euler and Gauss were not referring to anything when they used numerals, and (barring a novel account of mathematical truth) this would imply the unpalatable conclusion that virtually nothing that they said in number theory was true. (1974, p. 214)

To spell this out, suppose upon reading *Word and Object* that I decide that from this day forth the only mathematical entities I shall commit myself to are sets. In particular, I arbitrarily decide that from now on "0" in my mouth denotes Ø, "1" the set {Ø}, "2" the set {Ø, {Ø}}, and so on. Thus, I am an objectual reductionist, of the context-independent stripe, who arbitrarily chooses the von Neumann reduction to avoid commitment to sui generis numbers. As explained, the multiple reducibility objection does not undermine my position. But now I am faced with a problem of how to interpret others: other mathematicians, past and present,

---

[15] Field points out that if "mass" is taken to be relativistic mass, the first of the next two sentences is true and the second false; if it is taken as rest mass, the first is false and the second true: (1) "Momentum is mass times velocity"; (2) "Mass is [frame of reference] invariant". He concludes that each of the two objectual reductions is inadequate and that a partial-denotation semantics for "mass" is required (1974, p. 208). Whatever we make of this argument in the case of "mass", it has no grip on our reductionism. Any acceptable set-theoretic reduction must respect the accepted truth-values of arithmetical statements. The only cases in which they differ are mixed-theory cases such as "0 is an element of 2". However, mixed statements of this kind play nothing like the role in mathematics that (1) and (2) play in Newtonian theory. They are classic "spoils to the victor" or peripheral cases a theory ought not to be judged on. An objectual reductionist semantics is, therefore, adequate to the central uses of arithmetical language (in pure and applied mathematics). Indeed, in the physical case, the problem is that each objectual reduction only captures half the story; each is a failure. In the mathematical case, in contrast, each objectual reduction tells the full story; each is a success. The multiplicity objection arises because of double success, not double failure.

and past selves (myself before reading *Word and Object*). My semantic decision cannot be binding on them—the meaning of *my* words *now* is to some degree up to me, but the meaning of *their* words is not—and there is nothing in their concept of number to prefer one set-theoretic interpretation over another. Thus, it seems that their arithmetical talk cannot be interpreted as referring to sets. It should instead be interpreted as intending to refer to *sui generis* numbers, and so by my lights it is founded on a mistaken presupposition. Depending on my theory of presupposition, then, I must construe their arithmetical statements as false or truth-valueless; either way, not true. But surely it is absurd to suppose that others' utterances of "1 + 1 = 2" are not true.

Now, this consequence does not seem to me fatal for a philosophy of mathematics. For example, an error theory that declares that our mathematical discourse presupposes platonism yet that platonism is false, with the result that "1 + 1 = 2" as uttered by most people though mathematically correct is literally false, does not seem to me to be ruled out of court. Indeed, such an error theory seems to be Field's own position in the later *Science Without Numbers* (1980). Moreover, if Field (1974) is right, a partial-denotation reduction overcomes the difficulty.

However, a stronger response is available. Suppose that, following my decision to adopt a certain set-theoretic semantics for arithmetical language, I interpret others—Euler, Gauss and so on—in the very same way. On this semantics, Gauss's statement of the quadratic reciprocity theorem comes out true. What are the objections to this? One could accuse me of putting words in Gauss's mouth. His proof was about prime numbers, not sets. But if I think that prime numbers just are sets, I am right to think that his proof was about sets after all. Speakers do not know all the properties of their subjects of discourse. When the ancient Greeks spoke about the sun, they spoke, unknowingly, about a hydrogen-helium star that generates its energy by nuclear fusion. The ancient Egyptians and Mayans were not only ignorant of the properties of the star at the centre of the Solar System, which they worshipped; they were also wrong about them. In general, it is not attractive to suppose that a speaker cannot refer to something if she lacks knowledge of some of its properties or is wrong about a few of them. For familiar reasons not worth rehearsing, this contention represents a wrong turn in semantics.

The objection thus trades on the fact that "about" in the locution "*S* is talking about *X*" has a direct and indirect sense. In the direct sense, the statement is true only when *X* is presented in a way that *S* herself is familiar with or indeed *S* herself has used (in the latter case, the indirect discourse represents the direct discourse *verbatim*). For example, this is the sense in which, if *S* utters "Prince Albert was married to Victoria" but is foreign to London, the report "*S* is talking about Prince Albert" would be true, although "S is talking about the man to whom the memorial in Kensington Gardens is dedicated" would be false. In the indirect sense, in contrast, we can truly say that *S* is speaking of *X*, even if *X* is a way of picking out X that S neither used nor was familiar with, as in the Kensington Gardens example. It is in this indirect sense that Gauss was talking about sets.[16]

---

[16] A related objection put to me is that psychologists and linguists might reject the reductionist interpretation as giving an incorrect account of the subjects' thoughts and language. But what reason would they have for doing so? The point of this section is to examine whether there are any such reasons. Observe in passing that the reductionist view is

Field's argument was that even if there is no first-person multiplicity problem—I can break the tie arbitrarily—there remains a third-person multiplicity problem. Yet, once my set-theoretic reduction is on the table, there is a genuine reason for preferring it to others, namely that it accords with my interlocutor's wish to mean by her number words what others, myself included, mean by them, and her wish for her arithmetic discourse to be interpreted as true.[17] Following the arbitrary choice of an explication for myself, a non-arbitrary reduction for others presents itself. Breaking the first-person symmetry also breaks the third-person symmetry.

Now, I don't deny that there are mathematicians who do not intend, explicitly or implicitly, to mean by their number talk whatever others mean by it. Nor do I deny that there are a few self-conscious arithmetical sui generists who cling to their interpretation so tenaciously that their arithmetical discourse cannot be interpreted according to whatever is the reductionist's understanding. These stridently non-deferential arithmeticians cannot be understood in this way. From the reductionist perspective, only an error theory is right for them, coupled with a pragmatic story distinguishing their mathematically correct utterances from the incorrect ones. But it should be obvious that very few of the billions of arithmeticians, amateur and professional, fall into these categories. For all but the most philosophically committed mathematical users, the intention to say something true (when uttering something mathematically correct) overrides any intention that their arithmetical discourse be interpreted along particular metaphysical lines. As long as the speaker has the intention to be interpreted truthfully, and this intention overrides any intentions to put a certain metaphysical spin on her arithmetical discourse, the reductionist is right to construe her as referring to sets. And as long as the intention to be interpreted as one's peers is the prevailing one, the reductionist is right to construe the speaker as referring to the sets the reductionist's utterances are about. That accounts for all *normal* users of arithmetic, that is, philosophically casual speakers who are willing to submit themselves to the arbitration of the linguistic community for the meanings of their words. And Field's argument was clearly based on the attractions of the thought that it is wrong to interpret a normal arithmetician as consistently uttering untruths.

This is not to say that no other interpretative constraints are in play. Euler's intentions are best described by saying that they are to express true statements and to mean what others mean by their number words when used in a sufficiently similar way to his own. But there are limits to what we can take him to mean. The eccentric interpretation that takes, say, a carrot to be the referent of "2" in Euler's mouth and a parsnip that of his "3" (and makes corresponding changes for the predicates, quantifiers and function symbols) is categorically ruled out, given others' usage. An eccentric reductionist of this kind has no choice but to be an error theorist about Euler's claims involving "2" and "3". But less eccentric reductionists need not interpret Euler's arithmetical claims error-theoretically and may respect his intended truth-values. What if several speakers in the community make different choices? What if I interpret numerals in the von Neumann way but you, as an equally convinced reductionist who decided to break the tie in another arbitrary way, interpret them as Zermelo ordinals?

---

compatible with speakers having distinctive arithmetical concepts (as mentioned in footnote 1) and (as is evident) having distinctive arithmetical vocabulary.

[17] "Arithmetical discourse" here includes pure mathematical statements such as "$3 + 5 = 8$" and applied statements such as "the number of people in this room is eight".

Since there is no longer a privileged interpretation, doesn't the third-person symmetry objection re-arise? It is not clear that there can be an exact tie from a given reductionist's perspective. The kind of situation envisaged is one in which A is a von Neumann reductionist and B a Zermelo reductionist, there are no other reductionists around, and A and B are identical in all other respects (none is more of an expert or more linguistically deferred to by the community, etc.). From a God's eye view (on the running reductionist assumption that there are no sui generis numbers), the situation is symmetric and there is no reason to gloss normal arithmeticians' utterances as the A-interpretation or B-interpretation. But now look at it from A's point of view: A wants to preserve as much agreement between her utterances and those of normal arithmeticians—in particular this is more important to her than preserving as much agreement between B's utterances and those of normal arithmeticians; she has chosen the von Neumann interpretation for her arithmetical utterances; if she chooses that same interpretation for others' utterances, she gets to respect the agreement constraint; and no other symmetry-breaking constraints are in play. It follows from all this that she should choose the von Neumann interpretation for others' arithmetical utterances. The key is that A, not being God, always has a reason to prefer her own interpretation in cases that are symmetric from God's point of view. The reason is her prevailing wish to mean by her number terms what most (normal) speakers mean by theirs.[18]

Of course, it's possible that A's recognition of the God's-eye-view symmetry of her interpretation with the B-interpretation could loosen her resolve to interpret normal speakers in her chosen way. She may well decide to adopt B's interpretation instead. Or, she may decide to consult with B so as to adopt a shared semantics. In general, if pressures of uniformity are sufficiently felt, reductionists might collectively opt for a particular reduction (perhaps for mathematical reasons). This could happen by fiat, as a coordinated response to this problem. But as explained, there is no inconsistency in A and B sticking to their guns in the given scenario, A maintaining the A-interpretation and B the B-interpretation.

Two points underlying our discussion bear emphasis. The first is that the proper semantics for someone's discourse may not fit all her dispositions. In interpreting a speaker's words to mean one thing rather than another, we look for the best fit, judged by certain criteria; this best fit may not be a perfect fit. Thus, it is a captious objection to a suggested interpretation of a speaker's words that it is not a perfect fit despite its respecting her prevailing semantic intentions. This point should be familiar not only from the philosophy of language but also from everyday experience of translation and interpretation. The second point is that most mathematicians, past and present, tend not to have marked opinions about whether their utterances are objectual or structural, or if objectual, whether they are about *sui generis* entities or not. Within certain constraints,[19] mathematicians are more or less happy to write a blank cheque as far as the interpretation of their mathematical utterances is concerned. In any case, what is beyond doubt is that philosophically highly committed mathematicians are few

---

[18] Observe that from A's point of view B does in fact utter truths—set-theoretic truths—when uttering arithmetical truths; they are simply different truths from the truths A utters with the homophonic sentences.

[19] For instance, one constraint might be that the interpretation is not about concrete objects; another that the interpretation is not too distant from the discourse's surface form; a third that the interpretation is systematic (which might be glossed as recursively generable).

and far between. This gives an interpreter greater licence in construing others' claims about numbers than in construing their claims about, say, tables and chairs.

3.2 Slippery slope objection

In a later phase, generally regarded as degenerate, Quine (1976) pushed his set-reductionism to apparent breaking point by insisting that everything— physical objects as well as numbers and other mathematical entities—should be reducible to set theory. For example, chairs are sets involved in the set-theoretic relation of being sat on by persons, who are also a kind of set. This omnivorous reductionism seems to be a *reductio* of set-reductionism—a parody of philosophy, even. Surely, if there's anything we believe, it is that we are not mathematical objects. Thus, the slippery slope objection—set reductionism about number theory slides into an absurd set reductionism about everything.

A proper assessment of this objection would involve a case by case evaluation of each reductionism's merits. We cannot undertake that here, so we content ourselves with two observations. First, no axiomatic economy is achieved by a reduction of the physical to set theory. Unlike the case of arithmetic and the rest of mathematics, the corresponding physical principles have to be imported into set theory rather than derived. And whether the reduction of the physical to set theory achieves an ideological economy is unclear since it depends on one's theory of properties. This illustrates the point that the assessment of reductionism has to be piecemeal. The economy achieved depends on the particulars of the case.

Second, omni-reductionism does not obey a general constraint on reduction. There is an entrenched semantic constraint that the interpretation of mathematical kind terms such as "set" and "number" cannot overlap with that of concrete kind terms such as "person" and "chair". To put it nonlinguistically, if philosophy is the best systematisation of *our* beliefs, there are constraints on how that systematisation proceeds. The systematisation that takes persons to be sets strays too far from our beliefs to pass as their systematisation. The contrast is with the case of "set" and "number", where no such semantic constraint is in place, or perhaps only a very attenuated one. There are no similarly entrenched principles.

In general, then, there are domain-specific reasons for or against reductionism, and one cannot naively leap from number-theoretic reductionism to omni-reductionism. Whether a given reductionism is acceptable or not depends on the balance of forces. That the proper balance points to reductionism about numbers and other mathematical objects but not reductionism about persons is prima facie consistent, as the case for person to set reductionism has yet to be made. In sum, yes, reductionist pressures, once given in to, do point towards omni-reductionism. But no, they might not take us to that unpalatable end point, because they might well be weaker than the pressures behind number-to-set reductionism, and because there are countervailing reasons against taking reductionism that far.[20]

---

[20] A related objection is that in light of the Löwenheim–Skolem theorems, reductionism leads to Pythagoreanism—the view that all of mathematics should be reduced to number theory (cf. Benacerraf 1998, fn. 28 p.57). However, this objection is unpersuasive. It assumes first-order axiomatisations of mathematical theories or non-standard semantics for higher-order axiomatisations, which arguably do not capture their mathematical content. And it also impoverishes mathematics: if all we have is arithmetic, then we cannot use analysis or set

3.3 Epistemological objection

The objection is simply that arithmetic cannot be reduced to set theory because the epistemology of arithmetic is more secure than that of set theory.

The response to this is straightforward: on the reductionist view, arithmetic is a branch of set theory but not the whole of it.[21] The epistemology of arithmetic is accordingly the epistemology of a branch of set theory, not that of set theory as a whole.

On a traditional conception of a foundation for mathematics, the foundation upon which mathematics is built is self-evidently acceptable to all rational beings. Reductionism need not go hand in hand with this kind of foundationalism: one can accept the former and reject the latter. Different reductionists have different stories to tell about what justifies set theory; yet none of them need buy into the view that set theory is self-evident or indeed more credible than the branches of mathematics it reduces. In fact, foundationalism may be turned on its head. Since all the branches of mathematics can be modelled in set theory, set theory inherits all their uncertainties. That ZFC entails the interpretation of branch B in set theory, B^s , means that any reasons to doubt B's consistency are inherited by ZFC. Hence, no logically omniscient subject can assign greater credibility to ZFC's consistency than to B's (assuming that $\text{Con}(B)$ and $\text{Con}(B^s)$ have the same degree of rational credibility and that the proof of ZFC's entailment of B^s is transparent, as it typically is). Conversely, good reason to believe in the consistency of set theory constitutes good reason to believe in the consistency of other mathematical theories.

In sum, set theory is the least likely branch of mathematics to be consistent. It does not claim to be more credible than the other branches of mathematics and it fails to meet the foundationalist ideal of self-evidence. That is compatible with its serving as the single theory into which all of mathematics is reduced. Reductionism is not foundationalism.

4 Meaning analysis versus explication

For each of the four reductionisms, the multiplicity problem, fatal for meaning analyses, is converted into a problem of terminological arbitrariness. I have explained that from the reductionist point of view this problem is trifling: given some labels and some objects, it is a question of how to assign the labels to the objects. It is rational for reductionists to choose some assignment, arbitrarily. This, I take it, was Quine's view of the matter, though to my knowledge he never responded to Benacerraf's article.[22] We then sketched why other objections to context-independent objectual reductionism, reductionism's classic form, do not succeed.

---

theory since these theories' consistency cannot be proved in arithmetic. Yet, we need set theory to show that first-order analysis has an arithmetical model.

[21] The equi-interpretatibility of ZF minus Infinity with Peano Arithmetic provides us with an exact measure of how much of standard set theory PA is equivalent to.

[22] Quine (1992) is a brief discussion of structuralism, which is concerned more with Lewis' structuralist reduction of set theory (or subsethood) to mereology (or mereology and the singleton function) and the relation of Quine's doctrine of ontological relativity to global structuralism than with Benacerraf (1965), which Quine does not mention.

It was not lost on Benacerraf that his arguments appear not to threaten the Quinean stance. In a retrospective article on "What Numbers Could Not Be", he wrote

> Even if the realistically driven reductionist is undermined by *l'embarras du choix*, not so with the holistic Occamite, who is not beholden to any notion of "getting it right" that transcends the best theory that survives ontic paring. That is what keeps Quine, Field, and their friends in business. (1998, p. 56)

However, as I understand Benacerraf, although he recognises that his arguments may not have undermined the Quinean stance, he nevertheless objects that that stance is not worth taking.[23] Quine is interested in what he calls an "explication" of arithmetic to set theory rather than a meaning analysis (1960, pp. 258–259). He tells us not what our arithmetic is about but what it should be: his philosophy of mathematics is prescriptive rather than hermeneutic. As long as set theory can do the job of arithmetic, one can dispense with commitment to sui generis numbers; what matters is serviceability. On Benacerraf's view, however, the philosophy of mathematics should be hermeneutic: it should tell us what *our* theories, with their current meanings, are about, and how we know them. Philosophy of mathematics is either about *our* mathematics—arithmetic, group theory, analysis, and so on, as uttered in mathematicians' mouths—or it is about nothing. Hence, a meaning analysis, "appears to be the only line of inquiry that seems at all sensitive to arithmetical practice—the context of use of the arithmetical expressions" (1998, p. 57); the analysis of number was "constrained as an exercise that aimed, following Frege, at identifying 'which objects the numbers really were' " (1998, p. 48); and, decisively, "our primary philosophical task is hermeneutic" (1998, p. 49 fn. 18). This explains why a key premise of the argument is the statement quoted in Section 1, now with added italics marking the appropriate emphasis, "any feature of an account that identifies a number with a set is a superfluous feature of the account (i.e. *not one that is grounded in our concept of number*)" (1998, p. 52). In short, Benacerraf attempts to wrong-foot reductionism by urging that though an explication may overcome the multiplicity problem, the philosophy of mathematics should not primarily concern itself with explications.

There is a quick response to this metaphilosophical objection: *our proposed reductions are in fact hermeneutic*. What goes for Euler and Gauss goes for other arithmeticians too: as explained, the objectual reductionist, say of the von Neumann stripe, may take other normal speakers of arithmetic to refer to the von Neumann finite ordinals. By definition, a normal speaker of arithmetic is anyone in the reductionist's linguistic community whose intentions to refer to whatever others refer to by their number terms in this community and to speak truthfully when uttering mathematically correct claims override his or her metaphysical views about the numbers (and nothing else undermines these intentions). As a matter of empirical fact, almost all speakers of arithmetic are normal. Thus, (objectual) reductionism as articulated here does *not* preach a think-with-the-learned-but-speak-with-the-vulgar sermon, recommending surface verbal agreement combined with semantic disagreement. It is compatible with the claim that reductionists mean the very same thing by their number discourse as normal arithmeticians mean by theirs.

---

[23] If it is intelligible in the first place, which Benacerraf hints that it may not be (1998, p. 57). He also imputes to the reductionist the view that it is "lacking any clear sense" to ask whether, even if sets can do the job of numbers, numbers might still exist (1998, p. 56). But he gives no reason for these strong claims.

To spell out this second line of defence, call our current total theory T^{actual}, and suppose for argument's sake that T^{actual} is different from the reductionist theory T^{red}, perhaps because our actual arithmetical discourse makes an indefeasible commitment to *sui generis* numbers. (These assumptions will apply throughout the rest of this section.) The hermeneutic constraint is then that a philosophy of mathematics that issues in a theory T is correct if and only if $T = T^{actual}$. Since $T^{red} \neq T^{actual}$, it follows that T^{red} is not an acceptable philosophy of mathematics.

The objection to this constraint is straightforward. If T^{red} is agreed to be theoretically superior to T^{actual}, surely it is irrational to refuse to replace the latter with the former. And as the reductionist sees it, T^{red} is indeed superior to T^{actual}. We may well be interested in the question of what various statements of our actual mathematics mean—that is, in determining what T^{actual} is—just as we might be interested in the question of what the various statements in phlogiston theory mean. But that does not imply that we should reject reductionism. If it is demonstrated to phlogiston theorists that oxidation theory is scientifically superior to phlogiston theory, they should accept the theory of oxidation.[24] Hence, it seems that the only way to engage with the reductionist is to argue that her theory is not the best one.

A direct challenge to T^{red}'s theoretical superiority might for instance challenge (global) ontological economy's claim to be a principle of theory choice or that of other reductionist motivations. Now, since my aim here is not to assess the motivations for reductionism but rather to explicate it and to answer the multiplicity objection against it, I shall set such questions to one side. Our interest is in investigating whether *if* one accepts the reductionist principles one should be worried by Benacerraf's multiplicity objection. The interest of this investigation naturally turns on there being something to be said for the reductionist principles, otherwise the present chapter would be an exercise in strengthening a corner of logical space no one should be interested in occupying. As a matter of fact, the majority of philosophers do find a criterion of ideological economy attractive, and would prefer, for example, a theory with a single primitive predicate to one with 513, even if the two theories are spatiotemporally equivalent. Mathematicians also tend to prize theories with fewer primitives, which points to T^{red} being mathematically superior to $T^{non\text{-}red}$ and arguably scientists do so too. Ditto for axiomatic economy, and perhaps ontological economy as well (though this is more controversial).[25] Moreover, it is a good question to ask those who dismiss economy principles whether they think there are any norms of inquiry other than empirical adequacy. If you take this line, it is hard to see why, for example, you should prefer the theory of relativity to the "theory" which simply lists the theory of relativity's empirical consequences. Though I have not attempted to appease those who see the reductionist motivations as muddled, I hope these necessarily brief remarks point to some of the difficulties associated with that perspective.

One way in which a fan of meaning-analysis-only might try to defeat the reduction-as-explication view is to say that any reductionist theory must be inferior to our actual non-reductionist theory because it lacks the theoretical virtue of expressing the statements of our

---

[24] I am assuming that phlogiston theory is not interpretable as oxidation theory. By "rejection" I mean theoretical rejection, not necessarily rejection in practical contexts.
[25] Of course, economy principles could be a facet of some deeper desideratum, for example explanatoriness.

actual arithmetic. The importance of this factor in theory choice is questionable—is conservatism that important a constraint?—but in any case the objection misconstrues the present dialectical context. The dialectical stage at which our discussion takes place is when it has been provisionally accepted that T^{red} (the theory that sets but not sui generis numbers exist) is theoretically superior to $T^{non\text{-}red}$ (the theory that both exist) and the multiplicity problem is then pressed. However, to argue that T^{red} cannot be judged superior to $T^{non\text{-}red}$ because $T^{non\text{-}red} = T^{actual}$ is to put into question the reductionist motivations by denying that T^{red} is superior to $T^{non\text{-}red}$ *modulo* the multiplicity problem. This denial is no part of Benacerraf's understanding of his own argument, or else "What Numbers Could Not Be" would have contained sustained arguments against principles of economy and other reductionist motivations, or arguments in favour of strong conservatism as a methodological principle.

A similar reply can be made to the objection that how good a theory is as a theory of numbers depends on its being a theory about *the numbers* rather than something else, and that T^{red} is a theory about sets but not *sui generis* numbers. This objection also misconstrues the present dialectical context. It has already been granted that a reductionist theory is superior to any non-reductionist theory *modulo* the multiplicity problem. After all, this problem is supposed to hit you if you think T^{red1} superior to $T^{non\text{-}red}$, T^{red2} superior to $T^{non\text{-}red}$, and so on; you then notice that no two of these reductionist theories can be true and, seeing no substantive reason to prefer any particular one, conclude that none is true. The multiplicity problem was never intended to apply to you if you did not think that T^{red1}, T^{red2}, and so on were better than $T^{non\text{-}red}$ in the first place. If you are already in that position, Benacerraf's argument is superfluous.

I have heard it said (compare one of the quotations from Benacerraf) that reductionism does not answer the question, "but what are numbers *really*?". This is just a mistake. Reductionism as here developed does in fact answer that question: it says that numbers really are sets. Or, to put it metalinguistically, that the numerals really denote sets. To many, there is a whiff of conventionalism about reductionism. Yet far from being conventionalist, reductionism as articulated here is a form of realism. Its conventional element lies only in the decision of which labels to use for various parts of our ontology.[26] Once we have settled on some principles of theory choice (say empirical adequacy, principles of economy, etc.), the ontology we should rationally accept is no longer up to us but is determined by these principles.[27] Compare an analogous question about the earlier twins example: "who is Olympia *really*?". The fact that her mother made an arbitrary choice about the denotation of "Olympia" at birth does not imply some sort of anti-realism about Olympia; the arbitrarily chosen baby *really* is Olympia.[28]

---

[26] Whether Quine, at least in some moods or periods, was conventionalist about ontology is another matter. To the extent that he was, we are not following him.

[27] The only way to charge reductionism of conventionalism is to maintain that principles of theory choice are conventionally chosen. But if you believe that, you are a conventionalist about inquiry full stop and there is nothing specifically pro- or anti-reductionist about your view.

[28] Of course, if realism is understood in such a narrow way that only *countenancing sui generis* numbers counts as arithmetical realism, then reductionism is not realist.

Finally, I turn to two considerations Benacerraf himself offers in passing against Quine (1998, p. 56). He first rhetorically asks what the standpoint is from which we judge our total theory. The answer is surely from the standpoint of our current theory (where else?). We make judgments of the relative superiority of two theories T_1 and T_2 (one of which may be our current theory) using our current theory. For example, if we accept a qualitative version of ontological economy, which enjoins us to minimise kinds, then if according to our current theory T_1 is committed to 159 kinds of things but T_2 to only 23, it follows that T_2 is preferable to T_1 in this respect. Likewise for other theoretical desiderata. We use our current theory to arbitrate between the potential total theories under our purview, and pick whichever it sanctions as best. How to weigh and aggregate various respects to return a final verdict is of course a tricky business, but theory choice is not a simple algorithmic process on anyone's view. That we are dealing with global rather than local theories may make the assessment more difficult, but I have not come across any cogent arguments to the effect that global theory choice is impossible, and Benacerraf himself does not adduce any such argument.

Benacerraf's second consideration is that two total theories could pass the tests equally well. What if theory T is just as theoretically virtuous as T^*? In that case, the multiplicity problem seems to re-arise, since there is no reason to prefer T over T^* or vice-versa.

A first response is that it is presumably rational to adopt one of the following strategies: arbitrarily pick one of the two total theories T and T^*; alternate between T and T^* (depending on context); give each of T and T^* 50% credence.[29] Arguably, the last response is correct. But there is no need to privilege any of these here: our claim is only that at least one of them is rational. For what other options are there? Accepting the next best theory in order to overcome the multiplicity problem is surely irrational. By assumption, there are two other theories superior to this third-best theory and it cannot be rational to adopt a knowingly worse theory in preference to a better one. Moreover, the alleged multiplicity problem could re-arise for the next best theories: perhaps there are two third-equal-best theories, and so on. Accepting no theory whatsoever in case of a tie at the top is also wrong; as with the previous response, it turns double success into failure.

Secondly, everyone faces this problem, whatever their philosophy of mathematics. In the absence of a guarantee that this situation could never arise in the empirical domain, the philosophy of science must in any case deal with this issue. Whatever the correct answer in the empirical case applies to that of a total theory. Historical attempts to claim that there could not be two equally successful yet different scientific theories—the most famous of which was the positivists' claim that two observationally equivalent theories would in fact be the same—have been unpromising.[30] If successful, such arguments would anyway presumably show that the problem does not arise for mathematics either.

Thirdly, the issue also arises for a meaning analysis. What if the two best meaning analyses of our current arithmetic conflict yet are equally supported? In other words, what if our

---

[29] Notice that these options are analogues of the first, second and fourth reductionist semantics canvassed earlier.

[30] The response Benacerraf parenthetically considers on his interlocutor's behalf is to question whether there could be theories with different contents that are as virtuous as one another (1998, p. 56). However, he does not explain why one might think that two equally virtuous theories must have the same content.

arithmetical discourse is indeterminate between two (or more) interpretations? There is no reason to think that the alleged problem cannot arise for meaning analyses.

Finally, we note that we are unlikely to be faced with this problem for long. History suggests that whenever two rival mathematical or scientific theories are roughly on a par, one of them eventually wins out. Indeed, it teaches that the pressure for uniqueness is such that if a tie is on the horizon, our standards are likely to evolve to ensure uniqueness. This is a contingent reassurance, of course, but soothing nonetheless.

I conclude that Benacerraf has not pointed to any compelling metaphilosophical grounds for rejecting explications, and, so long as philosophy is primarily in the business of coming up with a best total theory, there do not seem to be any such grounds. And incidentally, as the reductionist sees it, the objectual reductionism articulated above is in fact hermeneutic, as explained. This completes reductionism's answer to "What Numbers Could Not Be". Reductionism is not damaged by the availability of incompatible reductions.[31]

References
Benacerraf, P. (1965), "What Numbers Could Not Be", *Philosophical Review* 74, pp. 47–73.
Benacerraf, P. (1998), "What Mathematical Truth Could Not Be", in M. Schirn (ed.), *Philosophy of Mathematics Today*, Oxford: Clarendon Press, pp. 9–59.
Field, H. (1974), "Quine and the Correspondence Theory", *Philosophical Review* 83, repr. (with Postscript) in his *Truth and the Absence of Fact*, Oxford: Clarendon Press, pp. 199–221.
Field, H. (1980), *Science Without Numbers*, Princeton: Princeton University Press.
Kitcher, P. (1978), "The Plight of the Platonist", *Noûs* 12, pp. 119–136.
Quine, W.V. (1960), *Word and Object*, Cambridge, MA: Harvard University Press.
Quine, W.V. (1976), "Whither Physical Objects?", *Boston Studies in the Philosophy of Science* 39, pp. 497–504.
Quine, W.V. (1992), "Structure and Nature", *Journal of Philosophy* 89, pp. 5–9.
Wetzel, L. (1989), "That Numbers Could Be Objects", *Philosophical Studies* 56, pp. 273–292