Abstraction and Additional Nature†

BOB HALE* AND CRISPIN WRIGHT**

In ‘What is wrong with abstraction’, Michael Potter and Peter Sullivan explain a further objection to the abstractionist programme in the foundations of mathematics which they first presented in their ‘Hale on Caesar’ and which they believe our discussion in The Reason’s Proper Study misunderstood. The aims of the present note are:

1. To get the character of this objection into sharper focus;
2. To explore further certain of the assumptions—primarily, about reference-fixing in mathematics, about certain putative limitations of abstractionist set theory, and about the effects of impredicativity in abstraction principles—which underlie it; and
3. To advance the debate of the issues thereby raised.

I

In ‘What is wrong with abstraction’, Michael Potter and Peter Sullivan [2005] further develop an original objection to the abstractionist programme in the foundations of mathematics. It is to the general effect that the use of abstraction principles implicitly to define fundamental mathematical concepts indefensibly assumes that the entities to which those concepts apply have no additional nature—no further essential properties—beyond what can be gleaned from the relevant abstraction principle. The abstractionist quite unjustifiably assumes, for instance, that there is nothing essential to cardinal numbers beyond what can be extrapolated from Hume’s Principle. This, they advise, was the point—to which, they allege, the present authors have been completely insensitive—of the analogy they drew in their earlier criticism of the abstractionist approach, between Hume’s Principle and the principle they called Members of Parliament:

† Thanks for helpful comments to Roy Cook and to an anonymous referee.
* Department of Philosophy, University of Sheffield, Sheffield S10 2TN, U.K. and Arché Research Centre, University of St Andrews R.Hale@sheffield.ac.uk
** Arché Research Centre, University of St Andrews, St Andrews KY16 9AL, U.K. and New York University cjgw@st-andrews.ac.uk
(MP) *a*’s Member of Parliament is the same as *b*’s Member of Parliament just in case *a* lives in the same constituency as *b*.

[Sullivan and Potter, 1997, p. 139]

Just as members of parliament have a nature going quite beyond anything that can be inferred from their identity-conditions as given by (MP) in terms of co-constituency relations, so—the suggestion is—cardinal numbers *may* have a further additional nature quite unreflected in *their* identity-conditions as given by Hume’s Principle. Whilst we do not accept\(^1\) that we did miss their point—at least as so far described—we have no alternative but to agree that we did not discern their strategic purpose in trying to make it. For we took it that their object was to press a version of the familiar Caesar problem—the obvious thought

\(^1\) As evidence of our persistent misunderstanding, Potter and Sullivan selectively quote from a footnote to our discussion of their earlier objection, which reads in full as follows:

So [Potter’s and Sullivan’s] thought is not that numbers may have an additional nature, undisclosed by Hume’s Principle, in the same way that MPs have an additional nature undisclosed by the principle *Members of Parliament*. That concern—in effect, that for all we have said, number *might not be a sortal but a functional concept*—is effectively precluded by the admission that Hume’s Principle is a necessary truth. [Hale and Wright, 2001, p. 395, fn. 83]

It is perhaps understandable that someone reading this passage in isolation—especially when the italicized words are omitted, as they were by Potter and Sullivan—should take us to have stubbornly missed their point. But that (mis)reading is clearly ruled out by the main text to which the note is appended:

It’s important to be clear about the limits of [Potter’s and Sullivan’s] particular objection. . . . It is being allowed . . . that Hume’s Principle is indeed a necessary truth, exactly as normally formulated, and hence that we know that it suffices for the existence of numbers that certain concepts enter into relations of one-one correspondence and that numbers are things which are essentially identified and distinguished from each other by reference to facts of one-one correspondence between concepts. What is being challenged is that anything has been done to ensure that this is the whole truth about them.

As any tolerably careful reader might have grasped, the purpose of the footnote was *not* to take back the interpretation of Potter’s and Sullivan’s point advanced in the main text, *but* to reject the analogy with *Members of Parliament*—to observe that if numbers do indeed have an additional nature, the way in which that comes about cannot be illuminated by comparison with the situation of *Members of Parliament*. The reason is exactly as stated: that while Hume’s Principle is acknowledged by both sides in the debate to be a necessary truth, *Members of Parliament* is a contingency (since a constituency can be unrepresented), and its contingency goes in train with the fact that what is characterised by the principle is not, as with genuine abstractions, the concept of a basic sort of entity, but a functional role which, in any particular case, nothing may occupy and whose occupants must, on that very account, have more to them than their discharge of that role. *Members of Parliament* is therefore a poor model for Potter’s and Sullivan’s purpose.
being that if cardinal numbers may possess an additional nature, what rules out that nature embracing their being people or, specifically, Zero’s additional nature being that of being Julius Caesar?—and we responded accordingly. But the critical thrust of ‘What is wrong?’ does indeed seem to be directed at making out that abstractionism suffers from another, quite distinct, weakness, unrelated to the Julius Caesar problem as that issue is normally (variously) understood. Getting clearer just what this distinct objection comes to, and explaining why we think it misdirected, will enable us to clarify some key features of the abstractionist approach which deserve more discussion than we have given them in earlier writings, and which seem not to have been well understood.

II

Is it really a possibility that the objects to which reference is introduced by an abstraction principle should have an additional nature—essential properties—of which the principle gives no inkling? Why might it be thought that abstraction principles could come explanatorily short, even in the best cases, in this kind of way. One line of thought, to which Potter and Sullivan seem attracted, sees the relationship between terms introduced by an abstraction—Hume’s Principle, say—and the objects for which they stand on a commonly accepted model of how reference is fixed for natural kind terms. They write:

What did Locke realise about ‘gold’? Effectively, that there is an element of blind pointing in our use of such a term, so that our aim outstrips our vision. Our conception fixes what (if anything) we are pointing at but cannot settle its nature: that is a matter of what’s out there. One image of the way [Hume’s Principle] is to secure a reference for its terms shares a great deal with this picture. [Sullivan and Potter,

2 In essence, our response was that within the framework for a treatment of the Caesar Problem proposed in The Reason’s Proper Study, the relevant form of possibility of ‘additional nature’, if it is to raise that problem, has to be that instances of a newly abstracted sortal may also be instances of some independent category of object—‘independent’ in the sense of being distinct from the category which maximally generalises the sortal concept in question. In The Reason’s Proper Study we argue that if this is a possibility, and is problematic, it is a problem that afflicts absolutely all our sortal concepts and has nothing especially to do with abstractionist foundations. William Stirton [2003] has recently called into question the adequacy of the line of solution there proposed. We defer discussion of his objections to another occasion.

3 Obviously they will typically have many contingent features which cannot be elicited from the abstraction principle alone, but require appeal to supplementary information. It is, for instance, just such a contingency that ‘\( Nx : x \) is a moon of Mars’ picks out an even number.
We ‘point blindly’, using some concept, or cluster of concepts, that apply at a relatively manifest or surface level in the hope of hitting off reference to something at an underlying level. It is indeed readily intelligible how the referent of an expression fixed in this manner might have characteristics—including essential characteristics—that are independent of the concepts by which the reference is fixed. Certainly if one took this to be an accurate picture of how Hume’s Principle ‘secures reference’ for number terms, one could hardly just assume that the conception of numbers which it imparts embodies the whole truth about their nature.

Potter and Sullivan offer little, however, to explain how this model might properly be transposed to the case of the abstract singular terms introduced by an abstraction principle. In fact, the Lockean model seems fraught with problems in the target case. For one thing, as they notice (whence their proviso: ‘if anything’), it goes with the model that it must be at least initially intelligible that a principle proposed in this spirit fails to hit off reference to anything. It cannot just be a given that reference is secured, even if it is—let alone that it is secured to entities of which the principle states a necessary truth. Rather, this is something which needs to be verified as a by-product of our, so to say, finding a range of objects ‘out there’ to which the conception embodied in the principle is (necessarily) faithful. Yet if that is to be possible, the objects in question must first be given to us under some other mode of presentation. What might that have been in the present case? The abstractness of the objects concerned precludes any form of demonstrative or ostensive presentation of them. And if it is suggested that they may instead have been given to us via some other set of principles they satisfy, the evident next question is how we might have assured ourselves that those principles effected reference to anything—for the Lockean model should presumably apply to them in turn. The issue is large and we have no space to pursue it here. But we record the strong suspicion that to try to understand the way Hume’s Principle secures a reference for its terms on the Lockean model is to enter an epistemological cul-de-sac—and challenge the opposition to show otherwise. However that may be, the fact is that the model of reference-fixing that Potter and Sullivan are proposing is simply a flat repudiation of the abstractionist conception of the matter. In briefest outline, the kind of contrast that is at stake is parallel, broadly, to the contrast between two ways of taking the question: given that a pair of things

---

4 We discuss it further in [Hale and Wright, forthcoming b].
both exemplify such-and-such characteristic manifest qualities—think, for example, of a list of the reference-fixers for ‘gold’ given in a way independent of any understanding of that term or an equivalent—do they have a property in common? Taken one way, the answer may be unobvious and negative: there may be ‘fool’s’ instances of a putative natural kind, or there may even just be no common kind underlying even normal cases of presentation of the qualities in question. Theorists who think of all properties in this way—sometimes termed ‘sparse’ theorists—will recognise a gap between a (complex) predicate’s being in good standing—its association with well-understood, feasible satisfaction conditions—and its hitting off a real worldly property. This conception stands in contrast with that of the more ‘abundant’ theorist, for whom the good standing, in that sense, of a predicate is already sufficient to ensure the existence of an associated property, a (perhaps complex) way things can be which the predicate serves to express. What Potter and Sullivan seem consistently to fail to appreciate is that the abstractionist metaphysics of objects, and reference to them, which they are so anxious to criticise, stands to the conception of the matter that they evince in their Lockean parallel as an abundant conception of properties stands to a sparse one. According to the abundant—‘neo-Fregean’—metaphysics of objects and singular reference, a justification for regarding a singular term as having objectual reference is provided just as soon as one has justification for regarding as true certain atomic statements in which it functions as a singular term. As with the abundant conception of properties, there is no additional gap to cross which requires ‘hitting off’ something on the other side by virtue of its fit with relevant specified conditions, as the property of being composed of the element with atomic number 79 is hit off (or so let us suppose) by the combination of conditions that control our unsophisticated use of ‘gold’. And just for that reason there is as little sense in the idea that objects of abstraction might have an additional essential nature as in the idea that an abundant property might essentially have more to it than is involved in its possessors’ satisfaction of the relevant (complex) predicate. If this conception of the matter can be upheld, there can be no legitimate concern that the reference of terms introduced by abstraction may be to objects certain aspects of whose essential nature is independent of the concepts in play in the relevant abstraction principles.

5 Of course the possessors may have more in common; they may indeed have an associated sparse property in common. But this will not be a matter of the relevant abundant property’s having an additional essential nature.
The debate, to this point, may seem pretty unsatisfactory. First it is so far obscure what our alleged misunderstanding consisted in. Second, it is anyway not clear—once the Caesar problem is set aside—why the mere (putative) possibility of additional nature should present a problem to programmes of abstractionist foundations, so long as it does not compromise our ability to know that the relevant abstraction principles are (necessarily) true—something which Potter and Sullivan have shown no inclination to question. Third, in re-iterating their insistence on the Lockean model of reference-fixing and the imagery of ‘blind pointing’, Potter and Sullivan have anyway simply gone past, rather than properly challenged, the conception of abstract objects and reference to them that is integral to our position, according to which, as just outlined, there is no clearly intelligible possibility of additional nature in the first place. What is going on?

The penny drops when one realises that when Potter and Sullivan emphasize the possibility of additional nature, what they have in mind concerns not the referents of abstract terms as introduced by successful abstraction principles—ones which perform as we (abstractionists) would like them to—but the, as it were, antecedent objects of the targeted mathematical theories, e.g., number theory, analysis and set-theory, which abstractionism aims to recover. Thus it is not their contention that, for example, the referents of numerical terms of the form, ‘$N x : A x$’, introduced by (finite) Hume’s Principle may have an additional nature, but that the natural numbers themselves may do so—and hence that abstractionist number theory may come short of capture of its intended objects. Their complaint, ironically, is not that terms introduced by successful abstraction may latch onto objects of unsuspected additional nature—for instance, Roman emperors—but, it appears, precisely its obverse: that just because the objects to which reference is introduced by a successful abstraction principle will not, at least in the intention of abstractionists, have any additional nature, but will be objects whose essence is, so to say, exhausted by the identity-conditions laid down in the relevant principle, that principle must fail to supply the basis for a satisfactory theory of any range of mathematical objects which do—or even may—have such an additional nature.

So that, then, was the misunderstanding. We took Potter and Sullivan to be claiming that abstracts—the referents of the terms introduced by successful abstraction—might yet have essential properties independent of the abstraction. But their claim is rather that mathematical objects as best (philosophically) conceived may do so, and hence that mere abstracts may be too ‘thin’ to play their intended role. The crossing of purposes was aided and abetted by the fact that ‘numbers’ equivocate in
the discussion, sometimes denoting the referents of the abstractive terms in Hume’s Principle, sometimes zero and its suite as pre-theoretically understood.

In fact, however, as Potter and Sullivan well realise, numbers—the finite cardinals—are a weak example for their purposes. They write,

> What we wish to claim about this case is only that, for all Hale and Wright have said, numbers may have a nature additional to that implied by their satisfaction of Hume’s Principle. We do not claim that numbers in fact have any such nature. Our point was only that the contention, for a given range of mathematical objects, that they have no such additional nature will require an argument specific to the case in question, and that Hale and Wright have neither supplied such an argument nor seen the need for it.

Given all this, the case of natural numbers is an unpromising one to use dialectically to persuade anyone of the possibility to which we are trying to draw attention [Potter and Sullivan, 2005, p. 190]

The suggestion in this passage that it is for an abstractionist account of some branch of mathematics to argue up front, as it were, that the relevant pre-theoretic objects have no additional nature exceeding what can be recovered from the account need not be taken terribly seriously. Provided that what look like the standard axioms have been derived from abstraction principles in good standing, and their meanings explained in a way that vouchsafes their normal applications, the abstractionist can perfectly decently wait for an opponent to produce an argument that something essential has nevertheless been left out. Potter and Sullivan evidently do not have the slightest idea how to mount such an argument for the ‘unpromising’ case of natural numbers. However they do think that they can do better for the case of sets.

Let us see.

IV

Potter’s and Sullivan’s anti-abstractionism about sets turns on their conception of the requirements of a satisfying resolution to the set-theoretic paradoxes. Such a resolution, they argue, needs to do more than draw merely on the distinctive identity-conditions for sets—since those are encapsulated in Frege’s Basic Law V, which, unchecked, leads straight to paradox. The extra, they suggest, that needs to be called upon is the essential  *metaphysical dependence* of a set upon its members: a dependence which ensures that no set can feature among its own
members, so that—as one corollary—there can be no universal set. Sets, they contend, are accordingly an example of a kind of mathematical object which do, as best conceived, have more to their nature than is implicit in their identity-conditions. Once again, the point is not that the objects to which reference is introduced by the abstraction principle(s) offered in some abstractionist set theory may have an additional nature. Rather it is just the opposite—precisely, that the objects, viz. sets, which the theory aims to capture do in fact have a nature transcending their identity-conditions, fixed by their essential relations of metaphysical dependence upon their members, and hence are not suited for abstractionist recovery. The problem is not that the abstracts may have an additional nature, but exactly that they cannot be presumed to have the actual (additional) nature of the intended ontology. We cannot resist pointing out that this objection is consistent with a complete repudiation of the Lockean model of the way in which the terms introduced by an abstraction principle acquire reference. It has no clear connection with—and indeed, is not well advertised by—its authors’ remarks about Locke and ‘blind pointing’. But it is, we grant, of independent interest. Potter’s and Sullivan’s idea is that, while nothing analogous need apply in other areas of mathematics, set theory’s troubled history imposes a special obligation on any foundational account of it: the obligation of providing ‘an explanation of, and solution to the paradoxes of set theory’ [Potter and Sullivan, 2005, p. 192]. In their opinion, no abstractionist account is going to accomplish this. For the best explanation and resolution of the paradoxes is simply nothing to do with the identity-conditions of sets but is grounded in their asymmetric dependence on their members. This receives its proper reflection in the iterative conception, whereby sets emerge as essentially well-founded and the pseudo-sets exploited in the paradoxes emerge as exactly that. Since any abstraction principle can only capture the identity-conditions of the relevant abstracts, abstractionism is doomed to miss out on that feature of sets which, once recognised, provides the best explanation of why the paradoxes don’t really arise but are merely creatures of a misconception; the feature, namely, which consists in their possession of the kind of ‘internal structure [that] the iterative conception seems to require’ [Potter and Sullivan, 2005, p. 191].

Although Potter and Sullivan disclaim any intention of arguing for any particular view of sets, their remarks belittling the prospects of abstractionist set-theory have little force unless one is willing to take seriously the idea that the iterative conception does indeed promise a kind of explanatory treatment of the paradoxes which, first, is needed and,

6 What is best explanation in this context? Objectively correct? Most satisfying? Most insightful?
second, holds out no prospect of being matched by abstractionist means. Their point, after all, is not merely to canvass the *sheer possibility in principle* that the best notion of set might elude abstractionist treatment, but the supposedly actually vivid possibility that the iterative conception of set might *be* that notion—and that, if so, it is beyond abstractionist recovery. We have three, connected observations to make about this.

First, on the content of the explanation being suggested. There are obviously major philosophical questions about the key notion of the ‘dependence’ of a set on its members—how exactly is it to be understood? why is it asymmetric? how is it supposed to be *known* that it is exemplified by sets and their members in the way that Potter and Sullivan believe? While it would take us too far afield to pursue these questions here, it does need emphasis that a degree of optimism about their treatment is crucial for Potter’s and Sullivan’s point, since it is on the possibility of satisfactory answers that the utility of the iterative conception for the purposes of the ‘additional nature’ objection to abstractionism depends. It is quite possible, after all, to view the iterative conception as instead simply a way of *finessing* the issues raised by the paradoxes—as a mere device for, so to say, paring away from the naïve universe of sets a sub-population that allows of an interesting coherent mathematical theory. Potter’s and Sullivan’s much more ambitious suggestion is that sets have a metaphysical nature that *enforces* the iterative conception as a comprehensive account of what sets there are. But they leave both the content of the dependency claim, and the underlying epistemology of sets and their nature, unclear enough to leave the professed explanatory advantages entirely unjustified. It is all very well to insist on the need to explain the paradoxes. Good explanations work with hypotheses of relatively clear content, which there is reason to take to be true. The first is not yet, in our view, the situation of the alleged metaphysical dependency of sets on their members.7 And as far as reasons to take to be true are concerned, Potter and Sullivan seem content to fall in with the aloofness from

---

7 Potter and Sullivan do acknowledge that the requisite notion of metaphysical dependency is in need of better explanation than it has hitherto received, referring to the more extended discussion of a range of proposals about it canvassed in Potter’s [2004]. The problem, however, is that while there are certainly some phenomena—for example, personal identity, or causation—whose reality one is reluctant to put in question merely on the basis of explanatory philosophical difficulties; it is otherwise with the postulated metaphysical dependency of sets on their members. The former are cases where a notion is already entrenched in ordinary thinking, and the explanatory philosophical difficulties characteristically take the form of paradoxes. But ordinary thought has no investment in the metaphysical dependency of a set on its members; and to the extent that the notion proves explanatorily recalcitrant, it is simply unclear in what we are being invited to believe.
epistemological issues regrettably characteristic of too many of the friends of pure set theory. It is, by contrast, a large point in the abstractionist approach that the central epistemological issues about the targeted theory—how the nature of its basic objects and their existence may be known about, and the verification of the basic laws about them—are central priorities from the start.

Second, on the alleged tension of abstractionism with the iterative conception, and its supposed inability to move beyond identity-conditions. In this context, Potter and Sullivan contrive to give the impression that abstractionist set theory is somehow intrinsically wedded to a sometime suggestion of George Boolos involving the conception of ‘limitation of size’. They set out the lines of battle as being between the iterative conception, with its allegedly proper emphasis on (or at least due reflection of) the metaphysical dependence of sets on their members, and the particular abstractionist conception which Boolos proposed and investigated, with its allegedly unfortunate emphasis on identity-conditions and limitation of size—a principle for believing which to be true, they allege, ‘the literature of set-theory is curiously free of arguments’. As is familiar, the principle Boolos investigated—what he termed New V—was an instance of the following schema:

\[(\forall F)(\forall G)(\text{Ext}(F) = \text{Ext}(G) \leftrightarrow [(\text{Bad}(F) \land \text{Bad}(G)) 
\lor (\forall x)(Fx \leftrightarrow Gx)])\]

—specifically, the principle that results when ‘Bad’ is interpreted as universe-sized. One immediate point to make about this is that the particular abstractionist strategy it illustrates is natural but by no means inevitable. The paradoxes teach us that not all concepts determine sets. A solution to them must therefore somehow provide for a distinction between those that do and those that do not, and corral the former together into a coherent mathematical treatment. Under the strategy announced by the schema, ‘Bad’ is a schematic characterisation of those that do not. So the more general project effectively proposed by Boolos’s approach is that of finding an optimal interpretation of ‘Bad’—one that both

8 Potter and Sullivan, 2005, p. 191. Russell’s [1906] considerations in favour of such a principle are dismissed as ‘little more than bluster’. In fact the idea that Russell there gestures at is in effect that Law V be withheld from application to concepts whose extensions are (or include something) not merely ordinals-sized but ordinals-shaped as well, conjecturing that this proposal captures the idea of what has come to be known, in a more recent terminology, as ‘indefinite extensibility’. Many have felt that the idea that indefinitely extensible concepts distinctively fail to determine sets is potentially insightful—certainly, it seems hardly more ‘blustering’ than our protagonists’ notion of the metaphysical dependence of sets on their members. For a more nuanced discussion of Russell’s idea, see [Hallett, 1984, pp. 183–185] and [Shapiro and Wright, 2006, passim].
incorporates insight into the generation of the paradoxes and maximises the proof-theoretic strength of the resulting theory. ‘Universe-sized’ is merely what Boolos himself actually suggested (it has the collateral advantage, in the context of a neo-logicist abstractionist project, of being second-order logically definable), and the proof-theoretic strengths and limitations of the resulting system are well known. But other suggestions are possible within the framework supplied by the schema—other interpretations of ‘Bad’—and indeed other approaches altogether besides instantiations of the schema. One may of course be pessimistic about these directions, as about any unfinished research programme. The point remains that there is no reason to associate abstractionism per se with limitation of size, still less to write off its prospects (which are indeed problematic, like all work on the foundations of set-theory) on the grounds of that association.

In any case, the fact is that there need actually be no abrasion between abstractionism and the conception that a targeted range of objects may have essential characteristics beyond those specified in a first approximation to a statement of their identity-conditions. Boolos’s approach precisely illustrates, for example, how abstractionism might accommodate the idea that it is essential to sets both to be individuated by their membership and to be ‘not too big’. A philosopher who thought that limitation of size was indeed an essential feature of sets, over and above their identity-conditions, would simply have been mistaken to conclude that this would put them beyond abstractionist account, since there is the option of appropriately enriching their identity-conditions to capture the point, exactly as Boolos illustrated.

That raises the question whether the spirit of the iterative conception itself might not in principle be available for incorporation into a version of Boolos’s New V schema in just the same way, perhaps with a characterisation of non-well-foundedness selected for ‘Bad’ and well-foundedness thereby incorporated into a more refined account of the identity-conditions for sets. Potter and Sullivan [2005, p. 191] anticipate this thought and observe that it would likely be stymied by problems concerning the ideology necessary to formulate the relevant restriction—for one would need to define a condition on concepts sufficient to ensure that the associated sets were well-founded, and it is obscure how that might be done without admitting occurrences of the relevant abstraction operator within the statement of the abstractive relation, with consequent compromise of the credentials of the resulting principle as a (non-circular) implicit definition (there would be set-terms

9 For details, see [Boolos, 1998, ch. 6], ‘Iteration again’, especially pp. 98-104.
10 For exploration of an instance of this direction, but one subject to the misgiving about to be aired, see [Cook, 2003].
figuring on the right-hand side of instances of the abstraction principle). Still, one might wonder whether it might not be possible to get the same result indirectly—whether, that is, there are not choices for ‘Bad’ which allow the relevant instance of the New V schema to be in good order by abstractionist lights and are such that the extensions of concepts which pass the test comprise only well-founded ones.

As it turns out, this is not so either. Jané and Uzquiano [2004] show that any purely logical instance of the New V schema—any instance in which ‘Bad’, and hence the abstractive relation, is definable using just the resources of higher-order logic—will allow of models in which the well-behaved abstracts (those corresponding to ‘good’ concepts) are not well-founded. But there is a third possible approach: the option of abstracting in the first instance to a wider class of entities, not yet conceived as sets, some of which may be allowably ‘ill-founded’ (in the sense of falling under concepts of which they are the associated abstracts, or falling under concepts of which the associated abstracts fall under concepts of which they—the original entities—are the associated abstracts, or ...) and then defining the sets proper as exactly those among this wider class of abstracts for which well-foundedness holds. Indeed, Boolos’s work already pointed up this possibility. We know from the Jané and Uzquiano result that his version of New V does not entail Foundation (Regularity) for his ‘subtensions’—the abstracts of concepts smaller than the universal concept. But say that a concept $F$ is closed just in case any abstract falls under it if all its members do. And say that an abstract is pure just in case it falls under every closed concept. Then Foundation holds for the pure abstracts issuing from Boolos’s abstraction.11 What is to stop the abstractionist with iterativist sympathies proposing that these should be viewed as the sets?12

Again, Potter and Sullivan seem to anticipate something like this kind of room for manoeuvre, writing that it

11 See [Boolos, 1998, ch. 6, pp. 100–101].
12 Of course such a theorist still has to deal with the well-known issues about recovering Infinity, Power (and perhaps Inaccessibility) if he is to offer up something recognisable as a sufficiently populous set-theoretic universe to be to the liking of most set-theorists. But Potter’s and Sullivan’s objection was that abstractionism must remain silent about the metaphysical nature of sets, not their extent.

One concessive point is appropriate, however. A major objective of the philosophy of set theory is indeed to find a conception of set which gives us some reason for thinking that a corresponding set theory would be consistent. One important attraction of the iterative conception of set is that it promises to do just that. If abstractionism proposes to found set theory by first setting forth a New V-style abstraction principle and then limiting attention to well-founded abstracts, their well-foundedness will not give us any reason for thinking that the more inclusive system is consistent. The issue of consistency will depend entirely on whether the original abstraction principle is consistent, and optimism about it will require independent grounds.
is easy to define in second order logic with a Class operator what it is for a class to be well-founded, and ... to show that the standard theory holds for such classes. [Potter and Sullivan, 2005, p. 192]^{13}

But then they continue

But this is not to adopt the iterative conception, nor, crucially, to make it available as an explanation for the non-existence of a universal class. ... [W]hat was wanted was an explanation of, and solution to, the paradoxes of set-theory; averting your eyes does not count as a solution.

However this seems confused. The kind of theory proposed would indeed provide an explanation of the non-existence of a universal set, for it would provide the means to show that the abstract it associated with the universal concept would not comply with the definition it offered of set, the motivation for which could be precisely that well-foundedness is required by the asymmetric dependency of sets on their members.

We have no space to pursue these issues further here. Suffice it to say that, so far as we can see, Potter and Sullivan have made no case for an essential tension between abstractionism and the iterative conception, nor have they disclosed any limits on the extent to which the abstractionism can provide hospitality to the metaphysical motivations, obscure though they presently remain, of the latter. Their pessimism about the issue seems to us as premature as their apparent confidence in the metaphysics of sets which they offer to the iterative conception.

Finally, on the demand for explanation itself. As remarked, the one clear lesson of the paradoxes is that not all concepts determine sets. One might take the view that there is then only a pragmatic issue—that of finding a way of marginalising the exceptions in a way that somehow conserves the spirit of the original incoherent form of comprehension for sets and leaves scope for a powerful theory with as much as possible of the mathematical interest that drove the original. Abstractionism can certainly accommodate this project: the strategy will be to try to refashion the concept of set, exactly as encapsulated in suitable descendants of Law V, to underwrite a theory of suitable strength. There may be a number of ways of getting interesting results, none with any but a pragmatic claim to be superior to alternatives. But this cannot be Potter’s and Sullivan’s perspective. If it were, their complaint about abstractionism’s oversight of the metaphysical nature of sets would make no sense. In order for that complaint to make sense, there has to have been something which the naïve theory, as encapsulated in Law V, overlooked—a respect in

^{13} They seem to use the terminology of classes and sets more or less interchangeably.
which it was false to the nature of sets, as determined independently of any conception of ours. Indeed, that much goes with the territory of the ‘blind pointing’ objection to abstractionism. For if sets may have a nature transcending a concept of them as fixed in abstraction principles, then it may also transcend the concept fixed in theories of another kind. To be sure, we know what Potter and Sullivan think the relevant additional nature is. But the matter to remark is that their conception of the need for a kind of explanation of the paradoxes which they reproach abstractionism for failing to supply seems to be wholly driven by, and dependent upon, an extreme realism about the nature of sets—a realism extreme enough to make sense of the project of trying to diagnose the respects in which naïve comprehension goes astray as, so to speak, a thesis of the natural history of sets. When they call for explanation, what they seek is to understand the principles of set existence not in the pragmatic spirit just adumbrated—or the abstractionist has no difficulty in joining the game—but in a realist spirit which seeks to winkle out the divine truth for which Law V fumbles and to arrive at what it is about the essential nature of sets which underwrites that truth, whatever it is. Abstractionism wants nothing to do with the kind of realism that creates this demand for ‘explanation’. Indeed it is the conviction that any such position is epistemologically hopeless, coupled with receptiveness to a ‘face-value’ construal of mathematical theory, that provides the single most influential motive for the abstractionist project in the first place. It is important that philosophers interested in these issues register the point that Potter’s and Sullivan’s insistence on the putative explanatory shortcomings of abstractionism in this area rests on a demand which, absent a full-blown platonist conception of the subject matter, should simply be rejected as misconceived.

V

There is a further central critical idea in Potter’s and Sullivan’s discussion, which is perhaps for our purposes the most important idea that it contains. This is a suggested connection between the Lockean conception of the way in which an abstraction principle succeeds (if it does) in attaching reference to its introduced singular terms and the impredicativity of the abstraction principles characteristically deployed in abstractionist foundations. Concerns about the impredicativity of the abstractions that have been proposed for foundational purposes—the fact that their objectual variables typically need to be understood as ranging over the very abstracts concerned if the principles are to have the requisite proof-theoretic power—have, of course, often been voiced and discussed in the literature. But Potter and Sullivan evidently think that they can give them a novel slant. In essence, their contention is
that impredicativity enforces the Lockean model, with the consequence that impredicative abstractions cannot have the epistemological status required by the foundational rôle in which we have cast them. However, before we can properly elucidate and address this concern, we need to clear away some misconceptions implicit in—or at least liable to be encouraged by—the taxonomy of ‘ways of taking’ abstraction principles that provides the starting point for their discussion. 14

There are, in Potter’s and Sullivan’s view, just three ways in which an abstraction principle may be interpreted, which they respectively label:

1. Predicative;
2. Free;
3. Impredicative.

They gloss the taxonomy as follows. Predicative abstractions, they write, are ‘entirely harmless’, but useless as a foundation for mathematics—at least, ‘they cannot deliver substantive mathematical theories such as classical arithmetic’. In free abstractions, next, ‘we leave it open . . . whether the terms on the left-hand side refer to anything. Whatever commitment the abstraction principle involves is then purely conditional; if there are such things as Σs, it says, these are their identity-conditions’. 15 Like predicative abstractions, free ones are again harmless but foundationally useless—‘to be of much use, [a free abstraction] will have to be afforded with an additional assumption that, at least in certain cases, the terms on its left-hand side do refer to objects’. But that additional existential assumption calls for independent justification, so that the abstraction principle cannot, by itself (even assuming a suitable background logic), serve to found any significant mathematical theory. Finally, impredicative abstraction principles, for their part, do ‘hold out the prospect of delivering as consequences significant mathematical theories such as arithmetic, analysis, or set theory’; but they are ‘philosophically problematic in ways that the first two are not’. Interpreting an abstraction impredicatively, they tell us, ‘involves assuming that the terms on the left-hand side do indeed refer and that what they refer to are objects falling within the range of the quantifiers on the right-hand side’.

14 [Potter and Sullivan, 2005, pp. 187-188]. All the quotations in this paragraph and the next are from this passage, comprising the first four paragraphs of their paper.

15 Strictly, of course, there are no (closed) terms on the left-hand side of an abstraction principle, whether we write it—as do Potter and Sullivan—as a schema: Σ(F) = Σ(G) ↔ F ≈ G, or as its universal closure with respect to F and G. What they mean is that it is left open whether the closed terms figuring in instances of the abstraction principle have reference.
This tripartite classification is seriously misleading in at least two respects. First, Potter’s and Sullivan’s decision to count an abstraction as impredicative only if it is assumed that its left-hand-side terms actually have reference (to objects lying within the range of the right-hand-side quantifiers) obscures the point that there is a perfectly good and standard sense in which an impredicative reading of an abstraction need involve no such assumption—one on which the right-hand-side quantifiers are classed as impredicative if they are so understood that the referents of the left-hand-side terms, if any, lie within their range. On this reading, impredicativity is entirely consistent with freedom.

Second, it is simply an error to claim, as Potter and Sullivan do, that if an abstraction principle is free—in the sense that it is not assumed that its left-hand-side terms have reference—its deductive progeny must be merely conditional, so that if it is to be of any use in deriving any substantial conclusions, it must be ‘afforced with an additional assumption’ that abstracts of the relevant sort exist. An abstraction’s being free of any assumption of reference on the part of its left-hand-side terms is not to be confused with its being (equivalent to a) conditional: ‘If there are Σs, then Σ(F) = Σ(G) ↔ F ≈ G’,16 nor should it be supposed that a free (unconditionalised) abstraction must be foundationally useless, because incapable of delivering interesting consequences unless combined with an assumption of the existence of Σs—on the contrary, the existence of Σs can be inferred from such an abstraction, given instances of its right-hand side as supplementary premises.

Both points are central to our conception of abstraction, on which abstraction principles are, in crucial cases, both free and impredicative. We do not assume the existence of referents for the associated abstract terms. Since the instances of an abstraction principle are material biconditionals, all that is required for their truth is that their left- and right-hand sides have the same truth-value, and this condition can be met by taking left-hand sides whose terms lack reference to be false—provided, of course, that their right-hand sides are likewise

16 That Hume’s principle should be so conditionalised is contended by Hartry Field (cf. [1989, p. 169]). But his reason for insisting upon conditionalisation is that he interprets unmodified abstraction principles so that their instances may be false as a result of having true right-hand but false left-hand components—the falsehood of the left-hand sides itself resulting from reference-failure on the part of their ingredient abstract singular terms. However, there is no need for conditionalisation if, as on the kind of free interpretation we favour, complex sentences incorporating empty singular terms may still be true, even when the atomic components in which those terms are immediately embedded may not be. For a more extended discussion of Field’s insistence on conditionalisation, and our reasons for resisting it, see [Hale and Wright, 2001, pp. 143–145].
false. If the abstraction is conceived as a stipulation, this coincidence in truth-value will be precisely what is stipulated, in effect. Since we do not assume that all well-formed singular terms in the language have reference, the underlying logic must indeed be free. This does not, however, mean that the inference to any instance of an abstraction can only be safely made when the abstraction is supplemented with an additional premise asserting the existence of referents for its left-hand-side terms. There will, to be sure, have to be restrictions on the quantifier rules—specifically, on universal elimination and existential generalisation. But these restrictions will be restrictions on the first-order quantifier rules. There is no necessity for any such restriction on second-order \( \forall E \). If reference failure for any complex singular terms resulting from second-order instantiation causes the inferred instance to be false, this will involve simply that the second-order generalisation from which it was inferred is itself also false—not that unrestricted second-order \( \forall E \) fails to preserve truth. That is, if something of the form, \( \Phi \Delta (A) \), is false because the application of an abstraction operator,

---

17 Actually, the situation is more complex than usually recognised. We are allowing for the possibility of empty terms, whether simple or complex. Suppose we take it that, at least in the case of atomic sentences, reference failure always results in falsehood. Then it is possible for \( (\forall x)A(x) \) to be true but an instance \( A(t/x) \) to be false. Hence we must have a restriction on \( \forall E \). The usual minimal restriction is that \( A(t/x) \) may be inferred from \( (\forall x)A(x) \) only if we have a supplementary premise asserting that \( t \) exists. This restriction can be formulated in various ways—the supplementary premise could be \( (\exists y)(t = y) \), or \( t = t \), or just \( B(t) \) for some atomic matrix \( B() \). However, if lack of reference for \( t \) does not invariably render a context \( A(t) \) false, this restriction will actually be more stringent than needed. Call a context \( A(t) \) reference-demanding with respect to \( t \) if \( A(t) \) cannot be true unless \( t \) refers. Then the minimal restriction on \( \forall E \) will call for the supplementary premise only when \( A(t/x) \) is reference-demanding.

This is relevant to the question of what restriction in needed on \( \exists I \). If \( A(t) \) is always false when \( t \) is empty, then there need be no restriction on \( \exists I \)—since there will then be no case in which \( t \) is empty, \( A(t) \) is true but \( (\exists x)A(x/t) \) is false. But if one takes some contexts \( A(t) \) to be non-reference-demanding, we may have \( A(t) \) true but \( (\exists x)A(x/t) \) false (because no object in the domain satisfies \( A(x) \)). In that case we must restrict \( \exists I \) as well, by requiring the supplementary premise in each case when \( A(t) \) is non-reference-demanding.

A number of critics, including especially [Shapiro and Weir, 2000], and [Rumfitt, 2003] have suspected a can of worms for abstractionism around the issue of balancing the need for freedom in the underlying logic with its possession of sufficient strength to subserve proofs of the existence of the requisite abstracts. We see no problem. Some of the points made in the text above are anticipated in our [2003] in response to Rumfitt. But an explicit and self-contained treatment of the issues is clearly desirable. We hope to offer this in future work.

18 At least there will be none provided it is granted that there is no question of failures of predicate—i.e., open sentence—reference in the language concerned. We set to one side the sort of Aristotelian qualms about the reference of uninstantiated predicates made much of in [Shapiro and Weir, 2000].
Δ, to the predicate, A, results in an empty singular term, then so is the generalisation, \((∀F)(ΦΔ(F))\). Thus such restrictions as would be involved in a free logic suitable for, e.g., abstractionist number-theory put no barrier—to take a crucial case—in the way of the inference, by second-order \(∀E\), of instances of Hume’s principle. In particular, we may still unproblematically establish the existence of 0, defined as \(Nx : x \neq x\), via the appropriate instance of HP together with the logical truth that \(x \neq x\) is self-equinumerous.

Our position, then, resembles Potter’s and Sullivan’s second, ‘free’ interpretation of abstraction principles in one respect—it is free of any assumption that the associated abstract terms refer, and does in consequence call for a certain form of free logic. But it differs from it in insisting that the legitimate instantiations of an abstraction principle are not restricted to mere conditionals to the effect that if there are such-and-such particular abstracts of the relevant kind, they are identical or otherwise according as a certain equivalence holds among other things of the relevant sort. Nor do the restrictions involved in a suitable free logic extend to the second-order rules. So for neither of those two reasons do abstraction principles require to be supplemented by additional existence assumptions before they can be utilised deductively as required. Since impredicativity is consistent with freedom, we can accept an impredicative abstraction as true without presupposing reference on the part of the relevant abstract terms. What, if anything, secures reference for those terms is no supplementary assumption, but the input of true instances of the right-hand side of the abstractive biconditional, which enforces the truth of the corresponding left-hand-side identities. Since these are atomic, they cannot be true unless their ingredient abstract singular terms have reference.

To summarise the message of this section: it is not true that, when we interpret an abstraction principle impredicatively, our options are either to presuppose reference (and so sacrifice freedom) or to take on an obligation to ‘search’ among objects given independently for some which comport themselves as the abstraction principle requires—so that reference is accomplished, if at all, in the manner suggested by the Lockean model, by successful ‘blind pointing’. If what we have said here is right, there is ample room for the combination of impredicativity and freedom involved in the view we have sought to uphold—so that, even in the impredicative case, a successful abstraction may provide a concept which gives us our only means of referring to the objects in question, and the success of this reference may, in good cases, be assured, without any need for further searching, by the truth of instances of the abstraction’s right-hand side.
It remains to see whether, having overlooked this—our actual view—in their threefold taxonomy, Potter and Sullivan nevertheless manage to do anything to close it off.

VI

Summarizing what they take to be the moral of their argument, Potter and Sullivan write:

There are two uses of abstraction principles that are clearly unobjectionable. If, first, we view Hume’s principle \ldots{} predicatively, it is harmless and plausible, but quite useless as a route to substantial mathematical theories. If, alternatively, we regard it as a truth about objects of which we already have a conception, then it can in appropriate cases be used, along with second-order logic, as the axiomatic base for a development of part of mathematics. But the epistemological crux will then be our explanation for its truth: this explanation will have to appeal to features of the objects in question. In the case of arithmetic, for instance, we need to explain how we come to know that Hume’s principle is true about numbers, conceived as antecedently grasped abstract objects. [Potter and Sullivan, 2005, p. 192]

Although they are not completely explicit on the point,\textsuperscript{19} it would seem from this that Potter and Sullivan would allow that when an abstraction principle is predicative—and so, they will insist, foundationally impotent—there is no objection to our thinking of it as introducing the concept of a kind of abstract objects which are, essentially, just as the abstraction presents them as being: there is no ‘hidden’ additional nature, and no more to them than can be extrapolated from their identity-conditions, as encapsulated in the abstraction’s right-hand side and whose existence may be assured by simple inference from right to left across instances of the principle. But not so, of course, in the only other case they allow to be ‘clearly unobjectionable’, in which the abstraction is impredicative, but in which we have an anterior conception of some objects among which the abstracts are to be found. In this case, any reference to those

\textsuperscript{19} In subsequent email correspondence (14 March 2005), Michael Potter is quite explicit:

If the quantification is predicative, then the objects introduced by an abstraction can be as thin as the principle implies, since we presume no more of them than that they are reflections of the legitimacy of the objectual way of speaking that the principle introduces.
objects must be understood in accordance with the Lockean model, and the truth of the abstraction principle must be a matter of its being true of objects of which we have an independent conception—so that there can be no question of our knowledge of its truth being a by-product of successful implicit definition, as the abstractionist proposes.

The passage quoted above may be intended to suggest, not only that no other view of how abstraction might work is clearly unobjectionable, but that any other view is clearly objectionable. But if so, that is just innuendo. What Potter and Sullivan need to argue—rather than merely insinuate—is that impredicativity enforces the Lockean model of reference-fixing for the abstract singular terms introduced via abstraction principles, and so leaves no space for the neo-Fregean account of how such principles can be known. What, on this point, is their argument? It not easy to say, since they neither explicitly acknowledge the need for an argument, nor address the issue directly. We shall, however, comment briefly on two lines of thought which might seem to be hovering nearby and which, if granted, might seem congenial to their claim.

First, it might be granted that when we construe an abstraction *predicatively*, it is indeed open to us to adopt a ‘thin’ conception of the abstracts—so that there is no more to them than can be extrapolated from their identity-conditions as given by the abstraction principle—but then contended that this is so only because it is then open to us to think of the abstracts as enjoying no existence independently of the process, as it were, of abstraction to them. When, on the other hand, we construe an abstraction *impredicatively*, there is no option but to think of the abstracts which are to lie within the range of the impredicative first-order quantifiers as having their being quite independently—‘in advance’—of the abstraction, and so as (at least potentially) having a nature which it at best only partially reflects.\(^{20}\) The thought, in other words, is that while it is open to us to adopt a kind of ‘creationist’ conception of the abstract objects introduced by predicative abstraction principles, and

\(^{20}\) There is no clear and explicit statement of this line of thought in [Potter and Sullivan, 2005]. But in the recent correspondence to which we have already referred, Michael Potter seems to us to come close to it. The passage from which we have already quoted continues:

If the quantification is impredicative, then the objects we speak about with the terms introduced by the principle must be supposed to be objects that ‘had their being’ in advance of the principle’s introduction—they are not mere reflections of the way of speaking the principle introduces.

There is certainly a strong suggestion here that if the quantification on the right-hand side of an abstraction principle is predicative, the objects ‘introduced’ by the abstraction need not—and perhaps should not—be viewed as existing independently, in advance of the introduction of the principle.
thereby to pre-empt the Lockean model, there can be no question of extending this conception to impredicative abstraction—for since the abstracts themselves already lie within the range of the right-hand-side quantifiers, that would require them to exist before they are created!

It would take us too far afield to evaluate this suggestion here, except to note that the last claim is not obviously correct, since a determined creationist might resist the suggestion that objects over which she impredicatively quantifies must already exist—to suppose so is, she may claim, to overlook the option of taking the domain of quantification to be unrestricted but always growing, as objects are created. But however that may be, creationism holds little attraction for us—indeed, we have gone out of our way, in earlier writings, to reject it—and plays no part in the motivation for our view in either predicative or impredicative cases. When an abstraction is put forward as an implicitly definitional stipulation, there is no attempt to create objects or stipulate their existence—what is created, if all goes well, is not objects, but grasp of a concept. Whether that concept has a non-empty extension then depends entirely upon whether or not instances of the abstraction principle have true right-hand sides—and that is not (generally speaking) a matter for stipulation, or creation. This point is entirely independent of whether the abstraction is predicative or otherwise. On the view we defend, the abstracts whose existence is assured by the truth of a predicative abstraction principle, together with the truth of the right-hand sides of some of its instances, ‘have their being in advance’ of our laying down the abstraction principle just as much as in the case of an impredicative abstraction. Thus if the mere existence of abstracts independently of abstraction imposed the ‘blind pointing’ conception, it would do so across the board, not just in the impredicative case.

In general it is, in our opinion, one thing to hold—as we do—that, where an acceptable abstraction principle has instances whose right-hand sides are true, the existence of the abstracts thereby required may be an objective, mind- and language-independent matter; and quite another to hold—which we do not—that the reference to them effected by the associated terms must be conceived on the Lockean model or, more specifically, that the abstracts concerned must be regarded as falling under some concept(ion) which we (could) already possess, independently and in advance of accepting the abstraction. There is no good reason to think that this distinction must somehow disappear just in virtue of an abstraction’s being impredicative.

A second line of argument that impredicativity imposes the Lockean model draws on the familiar idea that quantification is determinate in meaning, or at least has determinate truth-conditions, only when the domain has been properly circumscribed. This requirement need not amount to a blanket prohibition on impredicative quantification. But in
the case that interests us—that of impredicative abstraction principles conceived as introductory of a sortal concept—what it will demand, if circularity is to be avoided, will be the possibility of a specification of the domain without recourse to the sortal our (impredicative) abstraction principle is being used to introduce. So the domain over which the impredicative quantifiers range will need to be given under the aegis of some independent conception of the objects belonging to it, among which any successful reference effected by terms introduced by the abstraction principle will then—or so it is plausible to suppose—have to be conceived à la Locke.

We do not know whether Potter and Sullivan would endorse something like this train of thought. But they do, near enough, anticipate our reply to it. For they acknowledge, in their closing paragraphs, that so long as we can legitimately take recourse to absolutely unrestricted (what they term ‘genuinely universal’) quantification, there will be no further difficulty in the idea that the first-order quantifiers in an impredicative abstraction range over independently existing objects of which we may have no antecedent conception—and indeed, we would wish to add, of which there may be no possible conception—indeed of the abstraction itself. But such recourse, they insist ‘...is highly questionable’. Their readers may find it obscure why they think so. The final paragraph of their discussion starts by canvassing a kind of unrestricted quantification—viz., quantification over ‘everything that our current conceptual scheme allows for’—which they apparently allow to be legitimate, but which, on one construal, is clearly unsuited for the abstractionist’s purposes (it would render the first-level quantifiers in Hume’s Principle, for instance, predicative. 21) They contrast this with what they seem to assume is the (more generous) sense of unrestricted quantification that we do intend: quantifying over ‘everything we may in the future find

21 At least it will do so if quantifying over ‘everything that our current conceptual scheme allows for’ is equated with quantifying over all and only those objects which fall under sortal concepts belonging to our current conceptual scheme, and if those concepts include none under which fall all those things of which we presently possess no specific concept, but which we may in future find reason to acknowledge. It is not, in fact, clear that the latter proviso is true. On the face of it, we can define a general sortal concept of object as follows: \( x \) is an object iff for some sortal concept \( F \), \( Fx \). The crucial question now concerns the range of the second-order quantifier—does it range just over the sortal concepts we actually possess? Or can it be taken to range over all possible sortal concepts? If the latter, then quantifying over everything our current conceptual scheme allows for really is absolutely unrestricted quantification after all. Even if that move is disallowed, our current conceptual scheme certainly does allow us to contemplate the possibility that our system of concepts may evolve to include concepts of which we as yet have no inkling, whose instances may fall under no sortal concept which we currently possess. So quantifying over ‘everything that our current conceptual scheme allows for’ can still be, in a perfectly good sense, quantifying over absolutely everything.
reason to include in our world view’. But that too is not what we intend. What we intend is *unrestricted* quantification, quantification over absolutely everything, whether we may in the future have reason to include it in our world view or not. Potter and Sullivan grudgingly concede that ‘it would be wrong, because inconsistent, to deny that we can’ quantify over everything in a way that is not restricted to objects of which we currently possess concepts or may in the future find reason to include in our world view. They claim, however, that

this negative point is hardly enough to justify ... the central presupposition of [Hale and Wright’s] approach: that there are undeniably some things we can say now about things undreamt of in our philosophy is surely too thin a basis for assuming that we have a grasp of full second-order logic as applying to them.

But what is their point here? Are they conceding that we can legitimately quantify over absolutely everything—all objects whatever, whether we have (future need of) concepts of them or not—or are they not? If they are, what further problem is supposed to attend taking the first-order quantifiers in full-second-order logic to be thus unrestricted? If they are not, what was it that they conceded in the sentence preceding that just quoted? Either way, they evidently have no argument to offer, else they would surely have given it.

The issue of the legitimacy of absolutely unrestricted quantification is, of course, much debated, and is scarcely to be resolved by a few quick remarks. Potter and Sullivan succeed in aligning themselves with a familiar position in that debate, but apparently have nothing to add in support of it.

VII

Let us take stock. We have an opposition between two broad conceptions of the manner in which reference may be secured for the terms introduced by an abstraction principle. The neo-Fregean conception has it that, in the best case, even if impredicative, such a principle may be laid down as true by way of successful implicit definition of a sortal concept, and that for those of its instances where the right-hand side holds true, the corresponding left-hand side will then afford us intelligible means of reference to an object for which we may have no other means of reference, nor any other concept, than that provided by the singular terms involved. The Lockean conception, by contrast, has it that there is no legitimate option but to understand the truth of the abstraction as answerable to the fitness of certain, so to say, antecedent objects (that is, objects for which independent means of reference, or anyway
some independent conception, is possible) to perform as required—to be correctly identified and distinguished in the fashion that the principle encodes. Since the existence of such a range of objects can be no matter for legitimate stipulation, such a principle cannot possess the epistemologically inaugural and minimal status neo-Fregeans propose for abstractions, but has to be recognised to be true in the manner of any other axiom proposed as apt for a previously given, or anyway independently determinable, species of mathematical object. Potter and Sullivan, as we understand them, hold a mixed view: the appropriateness of the two opposed conceptions lines up with the distinction between predicative and impredicative abstractions respectively. But how to assess this view? Neither of the two lines of thought discussed in the previous section offers anything to enforce it. Potter and Sullivan themselves provide no other argument. But nor have we offered any direct argument to reconcile impredicativity with the neo-Fregean conception.

Actually, it is difficult to ‘reconcile’ theses for whose tension one with another no even prima facie convincing case has been made. But we offer the following argument. Take Frege’s example of the abstraction for direction as standardly presented,

\[(\forall a)(\forall b)Da = Db \iff a \parallel b,\]

where ‘a’ and ‘b’ range over lines and directions are taken to be distinct from lines. This is a predicative abstraction for which, so Potter and Sullivan seem content to grant, the neo-Fregean conception of reference-fixing can stand. But suppose we ‘impredicativise’ the principle thus:

\[(\forall x)(\forall y)Dx = Dy \iff x \parallel y,\]

where it is now to be understood that the range of ‘x’ and ‘y’ is unrestricted, including in particular directions themselves. Obviously the original predicative version of the principle is among the consequences of this. But a new question is now raised about iterations of the D-operator—about how to evaluate the right-hand side when x and/or y take a direction as their value. There are no doubt several things we might do in response. For the case where both x and y are directions, we might stipulate, for example, that they are parallel if and only if they are (ancestrally) of parallel lines; or that all directions are parallel; or that none are. For the case where only one of x and y is a direction, we could stipulate that no direction is parallel to any non-direction; or that a direction is parallel to any line parallel to a line of which it is (ancestrally) the direction; or a number of other things. There may be advantages and disadvantages among these various moves, but there would be little difficulty in making them give consistent results while preserving the entailment of the original predicative principle.
importantly of all, we could just leave the whole issue undefined. In short, it seems to be of no consequence what we do. The consequential content of the impredicative abstraction seems to be exactly that of the predicative original. But if Potter and Sullivan were right, there would have to be to be more: the impredicativity ought to have quite another, additional and highly significant effect, viz., that one who clear-headedly grasps the impredicative principle will understand it as holding out a substantial hostage, since the question of whether there are any objects suitable to serve as the reference of the \( D \)-terms now needs to be settled, if indeed it can be settled, by a kind of independent check over the domain. But why should that be so? And how might we be made to see it? Why does impredicativity, rather than merely opening up the questions about iteration and the content of ascriptions of parallelism to directions, force us into regarding the issue of reference as now unsettleable by the resources provided by the abstraction itself and the behaviour of the abstractive relation? Why, to the contrary, cannot the objects it calls for just be the same objects as those whose existence may be seen as settled—or so Potter and Sullivan seem content to allow—by the resources provided by laying down the first-order abstraction which the impredicative principle entails and by the behaviour of its abstractive relation?

Our position remains, unrepentantly, that even in the impredicative case, the newly introduced terms may provide our sole underived means of reference to the objects in question, and that warrant to regard those terms as referring need be sought no further afield than in the abstraction principle itself, along with our ordinary and antecedently available means of appraising the truth of instances of its right-hand side.\(^2\)

\(^2\) Charles Parsons [1997] bruits a distinction which it is worth briefly contrasting with that proposed by Potter and Sullivan. Speaking of the idea that abstractions provide means for a ‘reconceptualisation’ of states of affairs consisting in the obtaining (or non-obtaining) of the abstractive relation, (e.g., parallelism) between a pair of entities in the abstractive domain (lines), Parsons proposes that with first-order abstractions, we may think of the reconceiving as simply a coarser way of individuating the objects in the domain and writes that

It seems the only thing there can be an argument about is whether to call the equivalence relation (parallelism) identity of a kind of objects

—adding ‘I don’t object to that’. However, he continues,

clearly the matter is different with second order abstractions where the equivalence relation is of concepts and the identities are said to obtain among objects. Boolos’s question, whence the mapping of concepts to objects, has force. [1997, pp. 270–271]

Parsons’s idea may be glossed as being that there is a clear case of an abstraction’s compliance with the neo-Fregean model of reference-fixing—or at least a case to which
References


______ [forthcoming a]: ‘Focus restored’, in Øystein Linnebo, ed., a special number of *Synthese* on Bad Company.


Shapiro, S., and A. Weir [2000]: “Neo-logicist” logic is not epistemically innocent’, *Philosophia Mathematica* (3) 8, 160–189.


He has no objection—namely where its effect is merely to propose a recount, so to say, of the objectual domain under the equivalence relation on its right-hand side. As he rightly observes, however, that idea cannot be extended to, say, Hume’s Principle, since there the abstractive relation is over a different—higher-order—range of entities. The contrast with Potter’s and Sullivan’s position is evident as soon as one reflects that Parsons’s point would be no less applicable in the case of predicative higher-order abstractions. And for our part, since we hold that the neo-Fregean conception may properly apply to second-order impredicative abstractions, we do not conceive it as grounded in or requiring the support of the kind of idea that Parsons canvasses for the first-order case (nor do we view that as a happy way of fleshing out the relevant idea of reconceptualisation, or ‘content recarving’; but we cannot pursue these issues here).