

## 8

# Logicism, Ontology, and the Epistemology of Second-Order Logic

*Richard Kimberly Heck*

## 1 What Is the Problem about Second-Order Logic?

At the basis of Frege's logicism is his logic. And, although Frege does clearly distinguish first- from second-order quantification, at least in *Grundgesetze*,<sup>1</sup> he never suggests that there might be any difference between them that is relevant to the logicist project. By contrast, even though Frege does insist that the ill-fated Basic Law V is a law of logic, he expresses reservations about its logical status even before Russell's discovery of the contradiction (Frege, 2013, p. vii; see Heck, 2012, pp. 32ff for discussion). Frege thus seems to have been as confident in the logical pedigree of second-order quantification as he was in that of first-order quantification.

Frege's philosophical heirs have, at least for the last half-century or so, tended to see things differently.<sup>2</sup> With his colorful accusation that higher-order logic is “set theory in sheep's clothing,” W. V. Quine (1970, p. 66) is no doubt the most familiar of Frege's opponents. But there have been others, too, whose reasons do not always align with Quine's. At least in the first edition of *Philosophy of Logic*, Quine puts a good deal of emphasis on what he calls the “staggering existential assumptions” that set theory makes, which he says “are cunningly hidden” in the formalism of second-order logic (Quine, 1970, p. 68). Quine's own concerns about second-order logic thus seem to have been rooted in ontology: no surprise, given Quine's obsession with matters ontological.<sup>3</sup>

<sup>1</sup> Though not in *Begriffsschrift*. See my “Formal Arithmetic before *Grundgesetze*” (Heck, 2016a, §3) for the historical details.

<sup>2</sup> It turns out to be none too obvious exactly how and why this happened. For some discussion, see Moore (1988), Shapiro (1991, ch. 2), Eklund (1996), and Ferreirós (2001). Thanks to the many people who responded to a question I asked about this history on the Foundations of Mathematics mailing list.

<sup>3</sup> Curiously, the cited passage is missing from the second edition, where it is replaced by a remark to the effect that the type restrictions inherent in second-order logic immunize it from paradox—though Quine does still remark, at the beginning of the next section, that “The set theorist's ontological excesses may

But there are other reasons to regard second-order ‘logic’ as unworthy of the name, and many of these concerns are broadly mathematical in character. We can divide them into two groups: those concerning the logical strength of second-order logic, and those concerning its expressive power.

To get a feel for the former, consider the first-order theory, in the language  $\{0, S\}$ , with the following axioms:

$$(S1) \quad \neg \exists x(0 = Sx)$$

$$(S2) \quad \forall x \forall y (Sx = Sy \rightarrow x = y).$$

This theory has a complete, axiomatizable, decidable extension (Enderton, 2001, §3.1).<sup>4</sup> By contrast, if we add full second-order logic to this theory, then we get a theory that is capable of interpreting second-order Peano arithmetic.<sup>5</sup> So, whereas the first-order theory of successor is mathematically trivial, the second-order theory of successor is very far from it, having sufficient strength to capture an enormous part (at least) of the theory of the reals. And the difference between these theories, note, lies solely in what is allegedly logic.

The easiest way to see the worry about expressive power is to note that there is a second-order formula that is valid (assuming the standard semantics) iff the continuum hypothesis is true (Shapiro, 1991, pp. 105–6), and one can cook up similar formulae for many other open questions in set theory. The *language* of second-order logic thus seems to allow us to express hypotheses that one might reasonably have regarded as not in the province of logic itself. Moreover, what logically implies what, in the second-order sense, will (assuming the standard semantics) equally depend upon how such set-theoretic questions are resolved: if CH is true, then the axioms of arithmetic will entail certain things; if CH is false, then they will entail other things.

Putting all this together, it is not hard to see why someone might think that, if your interest is in the question what the logical consequences of certain assumptions are, then second-order logic (assuming the standard semantics) is the wrong tool for the job.<sup>6</sup> The worry is simply that so much mathematics has been *built into* the ‘logic’ that the ‘logical’ consequences of a given sentence end up including a

sometimes escape public notice... disguised as logic” (Quine 1970, 1986, p. 68). I would guess that the change is due to Quine’s having compared comprehension to naïve abstraction just a few sentences before, which might have misled some readers. That he had to replace (rather than augment) the remark about “staggering existential assumptions” was presumably due to a need not to upset the pagination.

<sup>4</sup> The extension in question is the result of adding all instances of induction for the language in question.

<sup>5</sup> By ‘full’ second-order logic, I mean second-order logic with unrestricted comprehension and with polyadic second-order quantifiers. The latter are essential, since the *monadic* second-order theory of successor is also decidable (Büchi, 1962). Unlike in the first-order case, we do not need axioms for addition and multiplication, since these can be defined in terms of the ancestral, as Dedekind (1902, §xi) shows.

<sup>6</sup> Perhaps the best recent source for this kind of argument is Koellner’s “Strong Logics of First and Second Order” (Koellner, 2010). Boolos (1998, pp. 302–3) registers a very similar complaint, which he more or less attributes to Gödel, though I do not know what source he had in mind. Feferman (1999) gives related arguments against a much stronger thesis, namely, that any notion that is permutation invariant is logical.

great deal that can be, and so ought to be, brought to light in the form of explicit premises or axioms.

It is at best unclear, however, what bearing these sorts of considerations have on logicism of any stripe, but especially Frege's. Logicism certainly can be understood as directed at ontological concerns, and the logicism of the positivists seems in part to have been so motivated (Benacerraf, 1995). But Frege's logicism and the form of neo-Fregean logicism made popular by Crispin Wright (1983) are primarily *epistemological* views: they are, in the first instance, views about the source of mathematical (at least arithmetical) knowledge. It is often emphasized that Frege needs the Basic Laws of his system to be laws of logic. But of course Frege also needs the *theorems* of the system to be laws of logic, which means that the inferences permitted in the system must be logical, at least in the sense that they preserve logicity. And something similar is true even when we back away from logicism strictly so called. If logicism is understood as the claim that the truths of arithmetic are analytic, then it is not enough to show that HP,<sup>7</sup> say, is analytic and that we can, somehow or other, derive the fundamental truths of arithmetic from HP. It needs also to be shown that, whatever those derivations involve, they preserve analyticity—or whatever other interesting epistemological property one thinks HP has.

If that is the right way to think of logicism, however, then we need to ask why second-order logic's impressive expressive power should even be relevant. It is easier to see why its logical strength might seem relevant. Perhaps the thought is that we are committing ourselves to a great deal when we blithely accept all the second-order consequences of whatever we otherwise accept—much more than we can possibly have a right to without significant epistemological cost. Perhaps a similar thought is behind the ontologically based objections: if second-order logic makes “staggering existential assumptions,” maybe we need to earn a right to them before we can earn a right to all the second-order consequences of our other beliefs. But such arguments have a worrying structure: since *A* somehow commits us to *B*, we can only have an epistemic right to *A* if we independently have an epistemic right to *B*.

That does not mean that the ontological worries do not need to be addressed. It does mean that we must always remember what the real issue is: the epistemological one. Consider, for example, these remarks, from one of Bob Hale's recent discussions of second-order logic:

[I]f one thinks of properties—the values of bound higher-order variables—as individuated purely extensionally, then, however much one emphasizes the supposed ontological differences between properties and sets, one will not be able to get away from the fact that properties behave just like sets in all logically important respects. Properties will be distinct just in case there is an object that . . . has one property but not the other . . . Quantifying over properties will be equivalent, in effect, to quantifying over all the subsets of the domain. But then one might as well use set theory, for the existential assumptions will indeed be much the same.

(Hale, 2013, p. 139)

<sup>7</sup> For those not yet in the know, HP says that the number of *F*s is the same as the number of *G*s iff there is a one-one correspondence between the *F*s and the *G*s.

Hale attributes this argument to Charles Parsons (1981, §VII), but it is clear enough that he means to endorse it. I am not sure, however, that this really is Parsons’s argument. I’ll return to the issue below. For the moment, let’s just discuss Hale’s version of the argument.

The obvious question is why the mere fact that “the existential assumptions” made by set theory and higher-order logic are “much the same” should imply that “one might as well use set theory” as use higher-order logic. After all, the existential assumptions made by Robinson arithmetic are *exactly* the same as the existential assumptions made by first-order Peano arithmetic—they are both committed exactly to the existence of the natural numbers—yet it would be absurd to think the two theories do not have different epistemological presuppositions. Why shouldn’t something similar be true of second-order logic and set theory? Might they too not have different epistemological presuppositions?

I am far from sure, then, why so many people have regarded such ontological concerns as central (or even relevant) to the question whether second-order ‘logic’ is rightly so called. Here, however, is a suggestion.

What we call second-order logic really comprises two components. First, there is an expansion of the *language* of first-order logic, through the introduction of second-order variables and quantifiers that bind them. We may think of these quantifiers as governed, minimally, by elementary logical principles corresponding to those that govern the first-order quantifiers, namely, in the case of the universal quantifier:

$$\begin{aligned} \dots F \dots &\vdash \forall F(\dots F \dots) \\ \forall F(\dots F \dots) &\vdash \dots G \dots \end{aligned}$$

subject to the obvious restrictions, and where  $G$  is required to be a variable. If that is all we do, however, then the result is a completely trivial extension of first-order logic. To get any logical power, we need to add so-called *comprehension axioms*,<sup>8</sup> which have the form:

$$\text{(COMP)} \quad \exists F \forall x [Fx \equiv \phi(x)].$$

Here,  $\phi(x)$  is some formula not containing  $F$  free, and there are similar principles for variables of other arities. We can think of each such axiom as saying that the relevant formula  $\phi(x)$  defines a ‘Concept’, that being the term I shall use, as neutrally as possible, for whatever is in the range of the second-order variables.<sup>9</sup> Which comprehension axioms we accept—which formulae we regard as defining Concepts—directly affects the power of the resulting system.

On one end of the scale, we might accept (COMP) only for formulae that contain no bound second-order variables (though they may contain free ones). This is known as *predicative* comprehension, and it is relatively unproblematic, since it amounts only to assuming that every *first-order* formula defines a Concept.<sup>10</sup> But predicative

<sup>8</sup> Equivalently, we can add substitution principles, or allow  $F$  not just to be a variable but also to be a formula. We will return to this matter below.

<sup>9</sup> I shall also speak of Concepts as ‘applying’ to objects, again trying to be as neutral as possible.

<sup>10</sup> In fact, it can be shown that adding predicative comprehension to a given first-order theory is equivalent (in one well-defined sense) to adding a weak theory of truth and satisfaction. Second-order

comprehension, even though it does not just yield a trivial extension of first-order logic, is still very weak.<sup>11</sup> The crucial point for our purposes is that predicative comprehension is insufficient for the proof of Frege's Theorem, that is, for the derivation of the fundamental arithmetical truths from HP. Øystein Linnebo (2004) has shown that predicative comprehension will not even allow us to prove the axioms of Robinson arithmetic from HP, at least not if we use Frege's definitions of the basic arithmetical notions.<sup>12</sup>

On the other end of the scale, we can accept (COMP) unrestrictedly. This is *full* comprehension, which is extremely powerful and more than adequate for the proof of Frege's Theorem. Full comprehension is much more problematic than predicative comprehension, however, precisely because it is not predicative but *impredicative*. Consider, for example, an instance of comprehension of this form:

$$\exists F \forall x [Fx \equiv \forall G \exists H (\dots G \dots H \dots x \dots)].$$

Here, our new concept  $F\xi$  is being defined by a formula that quantifies over all the Concepts there are—including, therefore, the very concept  $F\xi$  that is being defined. If we accept *this* sort of comprehension, then, we cannot take the second-order domain to be comprised only of Concepts that are *first-order* definable, and it would be incoherent to take it to be comprised of all those Concepts that are *second-order* definable.<sup>13</sup> It makes no sense to talk about what Concept a formula like  $\forall G \exists H (\dots G \dots H \dots x \dots)$ <sup>14</sup> defines until it has been said what is in the range of the second-order variables: that very same formula may define different Concepts depending upon what the second-order domain is.

The crucial question about second-order logic, so far as neo-Fregean logicism is concerned, is thus *how impredicative comprehension is to be justified*. More precisely, the question is whether our entitlement to the instances of comprehension needed for the proof of Frege's Theorem can be shown to be free of significant epistemological presupposition. And here is where the source of the concerns about the relationship between second-order logic and set theory finally becomes apparent. Why, after all, do we think that impredicative comprehension is valid? What we need to know is that, whatever Concept the formula  $\phi(x)$  might define, that Concept will be in the second-order domain. Since  $\phi(x)$  presumably defines some subset of the first-order domain, it will be enough if, for every subset of the first-order domain, there is a Concept in the second-order domain that applies to exactly its members.

quantification can then be simulated by quantification over formulae. This idea seems first to have been explored by Parsons (1974) but has been explored in much more detail recently by Visser (2009).

<sup>11</sup> Adding predicative comprehension to a given first-order theory always produces a conservative extension of that theory. The resulting theory is still logically stronger, however, in the sense that it is not interpretable in the original theory (Visser, 2009).

<sup>12</sup> We can use other definitions to get Robinson arithmetic, as Burgess (2005) shows, but no one, I think, would regard those definitions as anything but gerrymandered (and Burgess does not suggest otherwise). We can also try ramifying (Heck, 2011c), or we can scale back our ambitions yet further and settle for an even weaker, but still non-trivial, theory (Heck, 2014). So all is not lost, even if we do think that only predicative comprehension is logical.

<sup>13</sup> At least, not if the first-order domain is infinite.

<sup>14</sup> I propose to omit a lot of quotation marks. No confusion should ensue.

That is, it will be true if the second-order domain is (or appropriately corresponds to) the powerset of the first-order domain. More precisely: if we take the second-order domain to correspond to the powerset of the first-order domain, then this (i) fixes the domain of the second-order quantifiers, (ii) thereby determines what Concept a formula like  $\forall G\exists H(\dots G \dots H \dots x \dots)$  defines, and (iii) guarantees that this Concept will be in the second-order domain, so that the relevant instance of (COMP) will be true. One might reasonably suspect, then, that the validity of impredicative comprehension *rests upon* the powerset axiom (and, indeed, upon other axioms of set theory, such as separation). If so, then our entitlement to impredicative comprehension presupposes an entitlement to the powerset axiom, and that will have obvious epistemological consequences.

In fact, however, we can generate a problem from even weaker assumptions. In the course of giving an argument mentioned earlier (see pp. 142–3), Parsons writes:

Consider the full second-order predicate calculus, in which we can define concepts by quantification over *all* concepts. If a formula is interpreted so that the first-order variables range over a class  $D$  of objects, then in interpreting the second-order variables we must assume a well-defined domain of concepts applying to objects in  $D$  which, if it is not literally the domain of *all* concepts over  $D$ , is comprehensive enough to be closed under quantification. Both formally and epistemologically, this presupposition is comparable to the assumption which gives rise to both the power and the difficulty of set theory, that the class of all subclasses of a given class exists. (Parsons, 1981, p. 166, emphasis original)

Parsons does not explain in what way impredicative comprehension and the powerset axiom are ‘comparable,’ but it should be clear enough that his worry is not that second-order logic and set theory are *ontologically* on a par but rather that they are *epistemologically* on a par. And his suggestion is not that we need to assume the powerset axiom in order to justify impredicative comprehension. It is, rather, that justifying impredicative comprehension is *relevantly similar* to justifying the powerset axiom: it is just as hard, perhaps, or requires similar resources. The close formal parallels between set theory and higher-order logic, to which Parsons also alludes, only serve to reinforce this point.<sup>15</sup>

In the end, then, I do not see why it should matter whether it turns out that, for every subset of the domain, there is a Concept that applies to exactly that set’s members. The crucial question is not whether the second-order quantifiers *do* range over the full powerset (or something similar). It is whether that assumption needs to be used in justifying comprehension or, more weakly, whether justifying impredicative comprehension all but amounts to justifying the powerset axiom.

I’ll return below to the question how I think impredicative comprehension should be justified. First, I want to discuss two recent attempts, by Wright and Hale, to free second-order logic from the “staggering existential commitments” with which Quine attempted to saddle it. As will already be clear, I think the focus of these efforts is

<sup>15</sup> It can be shown that  $n^{\text{th}}$ -order arithmetic is mutually interpretable with Zermelo set theory with a restricted version of the powerset axiom saying just that  $\mathcal{P}^{n-1}(\mathbb{N})$  exists. As a special case, then: second-order arithmetic is mutually interpretable with Zermelo set theory with just the axiom that  $\mathcal{P}(\mathbb{N})$  exists. McLarty (2013) gives a rigorous account of the details.

misplaced. But the more important point will be that neither Wright nor Hale pays enough attention to what I have just argued is the really important question: how impredicative comprehension can be justified and what its epistemological presuppositions might be.

## 2 Wright on ‘Neutralist’ Quantification

In “On Quantifying into Predicate Position,”<sup>16</sup> Wright ventures to free second-order logic of all existential commitment by articulating what he calls a ‘neutralist’ conception of quantification. This view has two distinguishable parts. First, following Agustín Rayo and Stephen Yablo (2001), Wright (2007, p. 153) suggests that we should regard a quantified statement as committed to no sort of entity to which its relevant instances are not already committed. This is not only a plausible thought but arguably a Fregean one (Dummett, 1981, pp. 223ff). As Hale puts it:<sup>17</sup>

The suggestion that “Someone wrote Hamlet” carries a commitment to the existence of people that is absent from “Shakespeare wrote Hamlet” makes a complete mystery of existential generalization, which is surely to be understood—as its name suggests—as *generalizing* a commitment, rather than *introducing* one. (Hale, 2015, p. 10, emphasis original)

As Hale goes on to note, however, issuing this corrective is compatible with retaining the familiar idea that quantifiers have a “range of values.” On that view, the thought would be that quantification allows us to speak generally about people by exploiting a prior recognition of their existence.

It is really this latter idea that Wright means to question. He wants to suggest that quantification into a particular position is intelligible independently of whether the expressions that would normally occupy that position refer to anything. Rather, Wright says:

[Q]uantification should be viewed as a device for *generalization of semantic role*. Given any syntactic category of which an instance, *s*, can significantly occur in a context of the form [... *s* ...], quantification through the place occupied by ‘*s*’ is to be thought of as a function which takes us from [... *s* ...], conceived purely as a content, to another content whose truth-conditions are given as satisfied just by a certain kind (and quantity) of distribution.

(Wright 2007, p. 154, emphasis original)

The following view then becomes available: predicates do not refer to anything, but that need not stop us from existentially quantifying into predicate positions; doing so introduces no existential commitment not previously present; so second-order logic has no distinctive existential commitments.

<sup>16</sup> Hale (2014, pp. 446–7, fn. 5) complains in his critical study of my book *Frege’s Theorem* that I do not discuss Wright’s paper. The reason is simple: although my book was published in 2011, the paper that takes up these issues, “A Logic for Frege’s Theorem” (Heck, 2011a), was essentially completed in 2006 and was submitted to the editors of the volume for which it was commissioned in 2007. That was before Wright’s paper appeared.

<sup>17</sup> Note, however, that the simple sentence “Shakespeare wrote *Hamlet*” is surely committed to no person other than Shakespeare. Our use of proper names therefore commits us to no more than the existence of the people for whom we have names. But it would be reasonable to regard the existential quantification as having a more extensive commitment: to people, named or otherwise.

This is a bold proposal, but, like many other bold proposals, it is difficult to believe. What is the phrase “generalization of semantic role” supposed to mean? Consider the following sentences:

John jumped near the bank.  
 John jumped off the bank.  
 John jumped to the bank.

Is Wright’s proposal meant to license the quantification “John jumped something the bank”? Of course, if the quantifier here were interpreted substitutionally, it would be perfectly legitimate, but that is not what Wright (2007, pp. 155–6) has in mind. Surely, though, prepositions do form some sort of ‘syntactic category,’ so the problem must be that we have no conception of what ‘semantic role’ the prepositions are playing here. If that is right, however, then quantification does at least depend upon there being some uniformity of ‘semantic role,’ and it seems but a short step from there to a notion of *semantic value* that will provide a range for the quantifiers.<sup>18</sup>

The more important question, though, is how the meanings of quantified statements are to be specified, if not in terms of a range of possible values. Wright considers two answers.

The first—the official one—is inferentialist: we are to fix the meanings of the quantifiers by specifying rules of inference that will constitute their meaning. Wright suggests, unsurprisingly, that the meaning of the universal quantifier should be fixed by the usual introduction and elimination rules. In the second-order case, these would be:

$$\begin{aligned} (\forall+) \quad & \dots \phi \dots \vdash \forall X(\dots X \dots) \\ (\forall-) \quad & \forall X(\dots X \dots) \vdash \dots \phi \dots \end{aligned}$$

Here, I have written “ $\phi$ ,” where Wright has “ $F$ ,” to emphasize that it is meant to be a *predicate* (open formula) rather than a variable. That “ $\phi$ ” rather than “ $F$ ” appears in  $(\forall+)$  gives rise to several complications<sup>19</sup> concerning how the usual condition, that “ $\phi$ ” not be free in any premise, should be formulated, but I shall not pursue that issue. The important issue concerns the appearance of “ $\phi$ ” in  $(\forall-)$ . We’ll discuss this below.

Though Wright discusses only the universal quantifier, it is easy enough to guess what he would say about the existential quantifier. But it has become almost a commonplace these days that “all” and “some” are just two members of a much larger category of ‘generalized’ quantifiers, which also includes such words as “most,” “few,” and “many.”<sup>20</sup> Surely what we want here is a general theory of quantification, not a special theory of two somewhat atypical quantifiers. But it is very unclear

<sup>18</sup> Compare Gareth Evans (1985) on the importance of what he calls “interpretational semantics.”

<sup>19</sup> And to one significant confusion. Wright (2007, p. 157) suggests that it is important that we allow universal generalization even when  $\phi$  is non-atomic, but I cannot see why that would be a good idea. One does not usually make such an allowance in formal systems.

<sup>20</sup> The formal study of generalized quantifiers was initiated by Mostowski (1957). For a good overview of the topic, see the article by Westerståhl (2015) in the *Stanford Encyclopedia of Philosophy*. The classic paper on the significance of generalized quantifiers for linguistic theory was written by Barwise and Cooper (1981).



how to extend the inferentialist approach to generalized quantifiers. What are the constitutive rules of inference that govern “most” and “few”? It is bad enough that the logic of these quantifiers has no complete axiomatization. But “most” and “few” are also context-dependent, and not just in the way “all” and “some” are. It isn’t just that context needs to provide an implicit domain (as in “Everyone is on the bus”). What counts as being “most” depends upon the context, and “most” is vague as well. So it is difficult to see what, other than a prior commitment to global inferentialism—the view that meaning is always fixed by constitutive inferences—could make one have any confidence at all in this sort of approach.

Wright (2007, p. 161) also offers a second answer to the question how the meanings of quantified sentences are fixed. He regards this as a mere ‘heuristic’ that supplements the inferentialist account, but I propose to consider it on its own terms for the moment. The idea was implicit in a passage quoted earlier:

[Q]uantification through the place occupied by ‘*s*’ is to be thought of as a function which takes us from [... *s* ...], conceived purely as a content, to another content whose truth-conditions are given as satisfied just by a certain kind (and quantity) of distribution of truth-values among contents of the original kind. (Wright, 2007, p. 154)

As Wright develops this idea, we are to think of the thought expressed by some sentence ... *s* ... as itself structured, with *s* contributing some relevant ‘thought-part’ to the thought expressed by the whole. The universal quantification  $\forall x(\dots x \dots)$  is then to be regarded as true just in case *every* thought of the form ... *t* ... is true (Wright, 2007, p. 162): that is, every thought is true in which the thought-part contributed by *s* is replaced by some other thought-part *t* of the same type.<sup>21</sup> One might think that this gives the game away, since now second-order variables will range over the thought-parts corresponding to predicates, but Wright (2007, pp. 162–3) argues that we need not interpret the view that way.

This account still has a difficulty with generalized quantifiers. Is “At least two *F*s are *G*” true just in case there are at least two thoughts of the form “*t* is *F* and *t* is *G*” that are true? Plainly not. The fact that both of these sentences are true:

- Mark Twain is a person and wrote *Huck Finn*.
- Samuel Clemens is a person and wrote *Huck Finn*.

does not imply that “At least two people wrote *Huck Finn*” is true. A similar problem will arise for “most,” “few,” and the like. It is not, of course, that this shows that the truth-values of the quantified sentences do not depend upon “a certain kind (and quantity) of distribution of truth-values among contents.” One could, in the meta-language, introduce a notion of equivalence between names (or whatever expressions are involved) and then characterize truth-conditions in terms of the resulting equivalence classes (cf. Heck, 2011d, §8.3). But that formal maneuver, though certainly available, only serves to highlight the fact that the fine-grained *contents* expressed by the instances are actually irrelevant to truth-conditions. What is

<sup>21</sup> It is not clear how Wright proposes to categorize thought-types, and one might wonder if that will be possible, absent some suitable notion of semantic value. But I’ll not pursue the point.

relevant is only what the names have in common, namely, their reference. If so, however, then the truth-value of “Most *F*s are *G*” depends only upon *which things* are *F* and *which things* are *G*, and a correct semantic theory should make that explicit, as standard treatments do: quantifiers, to first approximation, express relations between extensions.

The objections we have been considering so far concern Wright’s ‘neutralist’ account of quantification in general. But even if these objections could be answered, there are other problems that arise if, as Wright intends, we want to use his account to answer questions about the role second-order logic plays in the neo-Fregean program. The reason is that, as Wright (2007, pp. 164–6) himself notes, his account has a difficult time justifying even very weak comprehension axioms.

As Wright formulates his introduction and elimination rules, there is no need for comprehension axioms. They are, in effect, built into the  $(\forall-)$  rule, by permitting the quantified variable in  $\forall X(\dots X \dots)$  to be instantiated by an arbitrary formula  $\phi$  (subject to the usual restrictions). But that does not make the problem go away: it just refocuses it on the question why we should regard the relevant instances of  $(\forall-)$  as logically valid. Since it is easier to discuss this matter with respect to explicit comprehension axioms, though, let’s do that.

The most obvious issue is with the impredicative case. Consider, for example, a comprehension axiom of this sort:

$$\exists F \forall x [Fx \equiv \forall G(\dots G \dots x \dots)].$$

On Wright’s ‘heuristic’ account, the formula on the right is supposed to be true of a given object  $x$  if all thought-predicates (as I shall put it) of the form  $\dots G \dots \xi \dots$  are true of  $x$ . The comprehension axiom, however, commits us to regarding the very thought-predicate expressed by  $\forall G(\dots G \dots \xi \dots)$  as being one of the thought-predicates that can replace  $G$  in  $\dots G \dots \xi \dots$ . The issue is not so much circularity as well-foundedness or, to borrow a term from Saul Kripke (1975), groundedness. In general, the evaluation of  $\forall G(\dots G \dots)$  depends upon the truth-values of all thoughts of the form  $\dots F \dots$ ; but the thoughts of that form will include thoughts of the form  $\forall G(\dots G \dots F \dots)$ . And it gets much worse, since there is no bound on how many second-order quantifiers, and what kind, might appear in such a thought-predicate.

There is a way out: we could restrict the thoughts on whose truth and falsity the truth-value of  $\forall G(\dots G \dots)$  depends to ones not involving second-order quantifiers. Then there is no danger of ungroundedness. But that will justify only predicative comprehension unless we can somehow argue that every impredicatively specified formula is equivalent to a predicatively specified one, i.e. unless we can somehow justify the infamous Axiom of Reducibility. Let me defer further discussion of impredicativity, however, since it will arise again in connection with Hale’s rather different account.<sup>22</sup>

Perhaps surprisingly, Wright’s account has difficulties already with predicative comprehension, as he himself notes (Wright, 2007, pp. 165–6). The problem arises with respect to instances of comprehension such as:

<sup>22</sup> See note 34 and the text to which it is attached.

$$\exists F \forall x (Fx \equiv Gx \vee x = a).$$

Here, both  $G$  and  $a$  are ‘parameters’: free variables in terms of whose values the new Concept  $F$  is being defined. The difficulty, for Wright’s view, is that  $F$  will exist only if there is some thought-predicate that has the described extension, and thought-predicates do not permit parameters. Formally speaking, there needs to be a formula in which *only*  $x$  is free and that has whatever extension  $Gx \vee x = a$  has, given whatever assignments are being made to these variables. As Wright (2007, p. 166) says, even for the case of *first-order* parameters, that is implausible if the first-order domain is uncountable, and it is all the less plausible if we allow second-order parameters.<sup>23</sup>

It is worth emphasizing how central to the neo-Fregean project comprehension with parameters is.<sup>24</sup> One place it is required is in the proof of a strong form of induction for the ancestral:

$$(IND) \quad Q^{*=}ab \rightarrow \forall F[Fa \wedge \forall x \forall y (Q^{*=}ax \wedge Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb]$$

This is needed for the proof of the existence of successors; it is also needed for the proof of arithmetical induction (Boolos and Heck, 2011, pp. 79–80). The crucial point is that (IND) is stronger than what we get directly from the definition of the ancestral:

$$Q^{*=}ab \stackrel{df}{\equiv} \forall F[Fa \wedge \forall x \forall y (Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb]$$

because of the presence of the additional conjunct  $Q^{*=}ax$ . The proof of the stronger form is not difficult: we just instantiate  $F\xi$  in the definition of the ancestral with  $Q^{*=}a\xi \wedge F\xi$  and appeal to  $Q^{*=}aa$  and a weak form of transitivity for the ancestral. But the comprehension axiom that is implicitly being used here

$$\exists G \forall x [Gx \equiv Q^{*=}ax \wedge Fx]$$

involves the parameters  $Q$ ,  $F$ , and  $a$ . The crucial case of (IND) needed for the proof of Frege’s Theorem is just:

$$P^{*=}0b \rightarrow \forall F[F0 \wedge \forall x \forall y (P^{*=}0x \wedge Fx \wedge Pxy \rightarrow Fy) \rightarrow Fb]$$

To prove that, we only need comprehension for  $P^{*=}0x \wedge Fx$ , but this still uses  $F$  as a parameter.

In fact, we already need parametric comprehension for the proof that predecession is one-one. Here’s the argument that  $Pab \wedge Pac \rightarrow b = c$ . Suppose  $Pab$  and  $Pbc$ . The definition of predecession is:

$$Pmn \stackrel{df}{\equiv} \exists F \exists y [n = Nx : Fx \wedge Fy \wedge m = Nx : (Fx \wedge x \neq y)].$$

<sup>23</sup> One might respond that, while we cannot have names for *all* of the uncountably many reals, we could have a name for *each* of them. I am not sure whether Wright would agree. But if we make that move in the first-order case, we can presumably make it in the second-order case, as well, and then it is unclear why the neutralist conception has any ontological advantage. We’ll return to this issue below.

<sup>24</sup> Formally, it can be shown that, if we limit comprehension just to the parameter-free cases, then—or so Harvey Friedman assures me—we cannot even prove such trivialities as:  $\exists F \forall x (Fx \equiv Gx \vee Hx)$ , i.e. that the second-order domain is closed under unions.

Since  $Pab$ , there are  $F$  and  $y$  such that  $b = Nx : Fx$  and  $Fy$  and  $a = Nx : (Fx \wedge x \neq y)$ ; since  $Pac$ , there are  $G$  and  $z$  such that  $c = Nx : Gx$  and  $Gz$  and  $a = Nx : (Gx \wedge x \neq z)$ . So  $Nx : (Fx \wedge x \neq y) = Nx : (Gx \wedge x \neq z)$ , and by HP, there is some relation  $R$  such that:

- (i)  $\forall t \forall u \forall v \forall w (Rtu \wedge Rvw \rightarrow y = v \equiv u = w)$
- (ii)  $\forall u (Fu \wedge u \neq y \rightarrow \exists w (Gw \wedge w \neq z \wedge Ruw))$
- (iii)  $\forall w (Gw \wedge w \neq z \rightarrow \exists u (Fu \wedge u \neq y \wedge Ruw))$ .

The rest of the proof involves our showing that

$$(*) \quad (R\xi\zeta \wedge \xi \neq y \wedge \zeta \neq z) \vee (\xi = y \wedge \zeta = z)$$

correlates the  $F$ s one-one with the  $G$ s, i.e. that it has properties corresponding to (i)–(iii). Then, by HP the other way,  $Nx : Fx = Nx : Gx$ , and so  $b = c$ .

For this argument to work, we need to know that (\*) can be used to instantiate the existential quantifier on the right-hand side of HP, i.e. that it defines a relation. The relevant instance of comprehension thus uses three parameters:  $R$ ,  $y$ , and  $z$ . And something similar is true of the other direction of the argument.<sup>25</sup>

It is tempting to suggest that, in the proof of Frege's Theorem, we can just assume that all the first-order parameters are natural numbers, which will all have names, in which case the relevant instances of comprehension will be available from Wright's position. But, first, the formal fact that we can take all the parameters to be natural numbers only falls out of the proof itself. One can hardly make that assumption going in. And this is not, in any event, how the neo-Fregean view is usually understood. Terms of the form " $Nx : Fx$ " are supposed to make sense no matter what sort of object the concept  $F$  applies to, not just in the mathematical case. One shudders to think how Frege would respond to the suggestion that we need to be able to name all the things to which  $F$  applies to be able to assign a cardinal number to that Concept (cf. Frege, 1980, §23). But that is the assumption we thereby make if we insist that the instances of comprehension used in the proof that precession is one-one must be free of parameters.

The really fundamental worry, however, is that nothing in the arguments we have just reviewed seems in any way to depend upon whether the objects appearing as first-order parameters are the sorts of things for which we do or even can have names. Any account of second-order quantification that makes it seem as if that could possibly matter is thus seriously flawed.

It would be an option, of course, to abandon the 'heuristic' and insist that the rules  $(\forall+)$  and  $(\forall-)$  fix the meaning of the second-order quantifier all by themselves. Since the latter permits impredicative instantiation, there can be no question about its validity. But there are different ways to take this suggestion. One would be to

<sup>25</sup> An even simpler example is the proof of:

$$\forall y \forall z (Nx : (x = y) = Nx : (x = z))$$

which is Frege's theorem 118. The proof consists in showing that  $\xi = y \wedge \zeta = z$  correlates  $\xi = y$  one-one with  $\zeta = z$ . For the proof to work, we need to know that  $\xi = y \wedge \zeta = z$  is a relation.

appeal to a strong form of inferentialism that regards such questions as whether the inference rules really do confer a determinate meaning on the quantifiers as illegitimate. I find this sort of view implausible, generally speaking, but the point is moot. Wright (2007, p. 155) agrees that any such account must recognize the legitimacy of *some* constraints on the rules posited, since some sets of rules lead to inconsistency, and he specifically suggests that the rules must satisfy conditions of conservativeness and harmony. But conservativeness is precisely what we do *not* get if we allow impredicative instantiation: quite generally, adding full second-order logic (or even  $\Pi_1^1$  second-order logic) to any consistent first-order theory (of reasonable expressive power) yields a non-conservative extension, since the consistency of the original theory will be provable in the expanded theory, as Tarski (1958a) famously showed.

I conclude that Wright's 'neutralist' account of quantification will not serve the needs of the neo-Fregean program.

### 3 Hale on Properties

Hale takes a different approach. He agrees with the usual view "that a quantified statement is properly understood only if it is known what the possible values of its bound variables are" (Hale, 2013, p. 154). But he argues that, if we accept a "deflationary" conception of properties on which "every meaningful predicate stands for a property or relation" (Hale, 2013, p. 133), then we may regard the values of second-order variables as being properties in this sense. If so, then there is no need to accept that the second-order domain contains a property corresponding to each subset of the first-order domain, and the Quinean worries about second-order logic's "staggering existential assumptions" may be dismissed.

Before we discuss this view, we need to get clear about its essential features. Hale tends to emphasize the intensional character of properties so individuated. For example, in a passage quoted earlier (see p. 142), he speaks as if the crucial issue is whether properties are "individuated purely extensionally" (Hale, 2013, p. 139). This is in a way peculiar, given Hale's focus on whether second-order logic is ontologically extravagant. Intensions outstrip extensions, and one might well wonder just how many intensions there are that correspond, say, to my singleton: there are ever so many ways in which someone might think of me. In the end, however, the question how properties are individuated turns out to be inessential. What really matters is whether "no more is required for the existence of a property than the existence of the objects in its extension" (Hale, 2013, p. 139, fn. 10). It is that sort of assumption that leads to the conclusion that, for every subset of the domain, there is a corresponding property in the second-order domain. Hale's view, by contrast, is that a property exists if, but only if, there could be some meaningful predicate that expressed it (Hale, 2013, p. 134).

As Hale notes, the view that second-order variables range over (deflationary, abundant) properties need not prevent us from using set-theoretic techniques to investigate second-order languages. It does, however, lead to a view on which the "possible values of the second-order variables [are] only sets which can be viewed as the extensions of genuine properties" (Hale, 2013, p. 145). Given how properties are being conceived, then, this suggests that the second-order variables will range "over all and only the *definable* subsets of the [first-order] domain" (Hale, 2013, p. 147, emphasis original).

What does “definable” mean here? We are never really told. It does not mean: definable in our language as it presently is. The view is that a property exists if there *could be* some meaningful predicate of some possible language that expressed it. But that does not help very much, absent some sense of what makes something a possible language. To be fair, though, Hale’s presentation of this view is fairly programmatic. He is more trying to make a certain view palatable than he is trying to establish it. Nonetheless, or so I shall be arguing, we know enough about the notion of definability and its relation to second-order logic to be justifiably skeptical that this view, any more than Wright’s, will serve a neo-Fregean’s purposes.

A first question is whether this view has any ontological advantages over the usual one. In his original paper on this topic, Hale (2013, pp. 145–6, 153) claims that his view will allow us to recognize only countably many properties, but Roy Cook (2014, pp. 5–6) shows that this position gets Hale in a fair bit of trouble. In a more recent paper, then, Hale is more careful:

Whilst any given finitely-based language can contain at most a countable infinity of predicates, so that at most countably infinitely many properties are definable in that language, it does not follow that the totality of all properties whatever . . . is at most countably infinite. (Hale, 2015, p. 6)

Even if all possible languages are finitary, that only implies that, in each language, only countably many properties will be expressible. But maybe there are uncountably many possible languages. If so, one might wonder whether Hale can even exclude the possibility that the totality of all properties should contain, for each subset of the first-order domain, a property true of exactly the members of that set. There need, after all, only be *some possible predicate in some possible language* that would express such a property, and the notion of possibility at work here is meant to be absolute metaphysical possibility (Hale, 2013, p. 135), which is pretty lenient.

Hale (2015, p. 7) suggests that, even if we assume that the first-order domain is countable, the utter arbitrariness of overwhelmingly most of the subsets of the first-order domain will prevent them from being the extensions of any possible predicate. But this is far from clear.<sup>26</sup> Let  $Q$  be some physical quantity that may, so far as physics is concerned, take any value in the real interval  $[0, 1]$ . The value of  $Q$  may be a length, a probability, whatever. Now let  $q$  be its value at, say, the exact moment of the vernal equinox in 2016<sup>27</sup> and consider the predicate:  $\xi$  is such that there is a 1 in the  $\xi^{\text{th}}$  position of the binary decimal expansion of  $q$ . For each subset  $S$  of  $\mathbb{N}$ , there is a real number in  $[0, 1]$  whose decimal expansion has a 1 in the  $n^{\text{th}}$  place iff  $n \in S$ .<sup>28</sup> So there seems no reason to think that the utter arbitrariness of a set implies that it cannot be the extension even of some *actual* predicate.

<sup>26</sup> Even if we do not follow Cook (2014, §3) and explore the consequences of allowing the languages themselves to be uncountable.

<sup>27</sup> I allow myself to assume, as seems reasonable in the present context, that we can specify at least one time sufficiently precisely that the quantity in question has some determinate value. Even if we cannot in fact do so, it appears to be enough, for reasons to be discussed below, if it is metaphysically possible for us to do so. Moreover, there seems no reason to suppose that our ability to specify the time precisely in any way constrains what value  $Q$  might take at that time.

<sup>28</sup> Take 1 itself to be represented as: 0.111 . . .

Moreover, it seems reasonable to suppose that there are examples of this form in which  $q$  might, at the moment in question, have taken on any value in  $[0, 1]$ . So, for any  $S \subseteq \mathbb{N}$ , there is a possible situation in which the predicate mentioned has  $S$  as its extension. It is unclear to me whether Hale would accept this as showing that (some property corresponding to)  $S$  should be regarded as among the values of the second-order variables. He speaks at one point of the second-order variables as ranging over “possible extensions of the properties for which . . . predicates stand” (Hale, 2015, p. 6). But this is importantly ambiguous. Is a set to be in the range of the second-order variables if there is some possible predicate whose extension that set *is*, in the actual world? Or is it sufficient if there is some possible predicate whose extension that set *would be* in some other world? Given his other commitments, Hale’s view must be the more permissive one. Surely we do not want *logic* to be sensitive to what values certain physical quantities happen to assume at certain times. In particular, if  $\exists F(\dots F \dots)$  is a statement of second-order arithmetic, its truth-value should not vary from one world to the next. But, for all that has so far been said, it could, if the second-order variables were to range only over sets that are the extensions of ‘possible predicates’ in some given world.

A second question is whether the ‘possible predicates’ are allowed to have parameters. As we saw earlier, parametric comprehension is essential to the neo-Fregean program. It is worth re-emphasizing this point here, however, since it is relevant to Hale’s claim to be able to prove the categoricity of second-order arithmetic, as he interprets it. The argument goes roughly as follows. Suppose that we are given two models of second-order PA. We then construct a certain map between them and prove that it is an isomorphism. In the course of that proof, we need to apply induction to formulae defined in terms of the domains of the models and the interpretations of the primitive symbols. Obviously, it is in the nature of the proof that the domains and the interpretations of the primitives are given by variables. Hence, the instances of comprehension needed in the proof of categoricity are parametric.

It has been shown by Cook (2014, Theorem 2.6) that, if we accept certain modal principles about the existence of possible expressions, then we can in fact recover full comprehension from parameter-free comprehension. But if we take that route, and restrict the second-order variables to definable subsets of the domain, then it is unclear that the proof of categoricity now works. Usually, one means by a ‘model’ a certain set-theoretic object: the domain is a set, and the interpretation of, say, “+” will be some set of ordered triples satisfying existence and uniqueness conditions. Hale’s proof of categoricity, however, cannot show that all models of second-order PA in *this* sense are isomorphic, since some (components of) such models may not be definable. We can, of course, speak instead of models that are characterized in terms of higher-order logic (see e.g. Rayo and Uzquiano, 1999), and then the proof will show that all models of this type are isomorphic. But one might be forgiven for wondering whether the proof only works because it uses a restricted notion of ‘model.’ Suppose we were to restrict ‘models’ of first-order PA to ones in which the domain and the interpretations of the primitives were recursive. Then it would follow from Tennenbaum’s Theorem<sup>29</sup> that first-order PA is denumerably

<sup>29</sup> Which tells us that there are no countable, recursive, non-standard models of PA.

categorical, i.e. that it has no non-standard countable models. Restricting what counts as a ‘model’ always makes categoricity easier to achieve—and completeness harder.

All of this is really so much skirmishing, however, since the main issue concerns the status of impredicative comprehension.<sup>30</sup> There has never been any question whether it is coherent to regard the second-order domain as restricted to *first-order* definable subsets of the first-order domain. But that gives rise only to predicative second-order logic. If we could say that the second-order domain comprised the *second-order* definable subsets, then that would justify impredicative comprehension. But we cannot say that. Definability is a semantic notion:<sup>31</sup> to say that a formula  $A(x)$  defines a subset  $S$  of the first-order domain is just to say that  $A(x)$  is true iff the value assigned to  $x$  is an element of  $S$ . It is therefore incoherent, because circular, to try simultaneously (i) to specify the second-order domain by saying that it consists of the ‘definable sets’ and also (ii) to allow second-order quantifiers to occur in the formulae by means of which the ‘definable sets’ are being defined.

The question is thus whether there is some other notion of ‘definable subset’ that will allow us to justify impredicative comprehension. But if we conceive of the second-order domain as consisting only of sets definable in certain terms, how can we be sure that we will not be able to define new sets *not* definable in those terms by quantifying over all the sets that *are* definable in those terms? Indeed, once the question has been put this way, one might well wonder whether some form of Grelling’s paradox does not loom: what are we to do about the predicate “predicate (or property) that does not correctly apply to itself”? Does it correctly apply to itself? The notion of definability must be handled with the utmost care. It looks very dangerous to assume that there is some totality of (extensions of) ‘possible predicates’ that is closed under quantification over that very totality.<sup>32</sup> It is in response to this problem that, in his work on definable sets of reals, Tarski (1958c) introduces a distinction between meta-language and object-language and insists upon a hierarchy of definable sets. That might suggest that there is no ‘definite totality’ of definable sets but only an ‘indefinitely extensible’ one. But, while that sort of move might save Hale’s account from paradox, it threatens to forfeit impredicative comprehension, since quantification over a given class of definable sets will always be able to take us outside that class.

I make these remarks only to illustrate how serious a problem impredicative comprehension poses for Hale, not to show that the problem is unsolvable. Hale (2015, pp. 13–14) is no doubt correct that, while taking the second-order domain to comprise (or correspond to) the powerset of the first-order domain will allow us to

<sup>30</sup> Some of the foregoing may also presume what Shapiro (this volume, §2) calls the “external” perspective. That said, if we assume the “internal” perspective, then, as Shapiro (p. 99) also notes, for every subset of the domain, there is trivially a property true of exactly that set’s members. Hale’s view will then need reformulating, at least.

<sup>31</sup> There is a corresponding syntactic notion, known by many names: numeral-wise expressibility, representability, and the like. So one might wonder whether that could be put to use instead. But it is a basic result that, if  $T$  is recursively axiomatized, then every set that is representable in  $T$  is recursive. Of course, this assumes that the logic is itself recursively axiomatized. But if it is not, then it is presumably given semantically, and the problem noted in the text now applies.

<sup>32</sup> This point is also mentioned by Cook (2014, pp. 9–10, fn. 17) and Shapiro (this volume, pp. 97–8, 108).



justify impredicative comprehension, it is far from obvious that there is no other way to do so. But Hale does not, it seems to me, really come to terms with this issue.<sup>33</sup> In his most recent discussion of it, he writes:

It is true that if we take as the range of the second-order variables just any proper subset of the power set of the first-order domain, it may well fail to be closed under definition by quantification. But . . . if the range is taken to be the set of all *definable* subsets of the first-order domain, definition by quantification over that range cannot take us out of it, since any subset so defined must, by its very nature, be one of the definable subsets. (Hale, 2015, p. 14)

The difficulty, as I have already said, is that, if we allow second-order quantifiers to appear in the formulae by which ‘definable sets’ are being defined, then we have *no satisfactory notion* of ‘definable set’ until the range of those quantifiers has been established. But if we do not allow second-order quantifiers to appear in those formulae, then it is not obvious how we are supposed to earn a right to impredicative comprehension.

Hale does not, however, really intend these remarks to justify impredicative comprehension. He says in a footnote that he is “assuming that impredicative specification is not to be ruled out on other grounds” (Hale, 2015, p. 14, fn. 25). But the question is not whether ‘impredicative specification’ is to be ruled out on *independent* grounds. The question is whether *Hale* is entitled to regard ‘impredicative specification’ as legitimate, given his particular views about what the second-order domain comprises. Hale is aware of the problem, writing in the same footnote:

[I]mpredicative specification of properties may be thought especially problematic when properties are conceived . . . as individuated by the satisfaction-conditions of possible predicates. More specifically, it may be feared that this will lead to situations in which the satisfaction-conditions for one first-level predicate,  $\phi$ , depend upon or include those of another,  $\psi$ , which in turn depend upon or include—with vicious circularity—those of  $\phi$ ? This issue is difficult, and I cannot discuss it properly here. (Hale, 2015, p. 14, fn. 25)

But this will not do. This is *the* fundamental issue facing any view of the sort Hale is defending (see also Shapiro, this volume, §4.1). One cannot just set it aside. To be fair, Hale does say a bit more:

I think the circularity need not be vicious. It is not difficult to find examples where impredicative quantification over properties in the definition of further properties or relations need raise no such problem. Pertinent examples are the Fregean definitions of precession . . . and its ancestral . . . which are unproblematic—at least provided that the bound property variables do not include within their range properties which can only be defined in terms of the precession relation. (Hale, 2015, pp. 14–15, fn. 25)

<sup>33</sup> Cook (2014, p. 2) argues that the deflationary conception of properties validates full comprehension, but he simply ignores impredicativity, instead focusing on the problem posed by parameters. Cook’s argument for comprehension just assumes, in the form of what he calls the Satisfaction Condition Existence Principles, that every predicate has a satisfaction condition. If this argument is to justify impredicative comprehension, then, Cook has to be assuming that predicates involving second-order quantifiers have satisfaction conditions. (See the very first steps of the proofs of Theorems 2.3 and 2.6.) But Hale’s entitlement to that claim is what I am questioning.

But the two cases Hale mentions are precisely where the problem we are discussing becomes most pressing.

It is worth emphasizing, first, that there is nothing impredicative about Frege's *definitions* of precession and of the ancestral. They can perfectly well be given, and used effectively, in predicative second-order logic (see e.g. Heck, 2011c). The issue concerns impredicative *applications* of these definitions: cases in which the bound second-order variables are instantiated by properties that are themselves defined in terms of second-order quantification. As mentioned earlier (see p. 150), this happens at a crucial stage in the proof of Frege's Theorem: in the proof of induction, we need to instantiate the universally quantified second-order variable that occurs in the definition of the ancestral with  $P^{*}=0\xi \wedge F\xi$ , which must therefore express a property in the range of that variable. Moreover, in the usual proof of the existence of successors, we need to instantiate the same variable with  $P(\xi, Nx : P^{*}=\xi x)$ , and we need to instantiate the existentially quantified second-order variable that occurs in the definition of precession with a formula of the form  $P^{*}=\xi a$ .

The problem is thus this.<sup>34</sup> Hale is certainly right that the required applications of Frege's definitions "are unproblematic . . . *provided that* the bound property variables do not include within their range properties which can only be defined in terms of the precession relation" and its ancestral (my emphasis). But the proof of Frege's Theorem *does* assume that the property expressed by  $P^{*}=0\xi$  is in the range of the bound second-order variables that occur in those definitions. That is precisely why impredicative comprehension is needed for the proof. If second-order logic, as Hale interprets it, is to be adequate for the proof of Frege's Theorem, then, we need to be given a reason, besides the fact that  $P^{*}=0\xi$  is syntactically a predicate, to think that the property it expresses is included in the range of the second-order variables that appear in its definition. Hale does not give us one.

There is an obvious way out, namely, to argue that the property expressed by  $P^{*}=0\xi$  can be defined otherwise than in terms of precession and its ancestral. But it would be utterly self-defeating for a neo-Fregean to insist that the property *being a natural number*—that being what  $P^{*}=0\xi$  is supposed to define—is included in the range of the second-order variables because it is definable in some other way. Which other way? In terms of finitude? This is exactly the classical worry—originally expressed by Bruno Kerry (1887, p. 295)—about Frege's definition of the ancestral.<sup>35</sup> Shall we instead define the concept of natural number in terms of set theory? Or physics? Or intuition? Any such answer will have serious epistemological consequences. The whole point of defining the concept of number in terms of second-order logic was to show thereby that it is a *logical* notion.

It is hard to escape a feeling that restricting the second-order domain to definable subsets will 'just work.' Take second-order arithmetic: isn't it obvious that we can allow the second-order variables to range over the subsets of  $\mathbb{N}$  that are definable in the very formal language in which the theory is stated? Once again, it is not: we do not know

<sup>34</sup> The brief remarks Wright (2007, pp. 169–70) makes about impredicativity are vulnerable to the same objections I am about to bring against Hale.

<sup>35</sup> My own discussion of Kerry's objections (Heck, 2016b, §2), and related ones, is focused somewhat differently from the present discussion, but it is also concerned with questions about impredicativity.

what sets of numbers are defined by such formulae until it has been said what the second-order variables range over; different sets will be definable given different such specifications. In fact, it is far from obvious that there is *any* range that might be specified for the second-order variables such that the sets that are definable, given that specification, are exactly the ones so specified as being in the range of the variables.<sup>36</sup>

It turns out that there is such a specification, but the construction involves some very heavy lifting. We can begin with the first-order definable subsets of  $\mathbb{N}$ . As said, if we take the second-order variables to range only over these, then we get just predicative second-order logic. But we can iterate the process. The idea of doing so is originally due to Bertrand Russell (1908), but in his hands it leads only to ramified type theory: even iterating the process through all the finite ordinals will not get us even a smidgen of impredicativity. Kurt Gödel (1938) would later realize, however, that the iteration could usefully be pushed into the transfinite. Gödel famously called the sets so defined the ‘constructible’ sets, which together comprise the ‘constructible universe’, nowadays known as  $L$ .

Let  $\Omega = L \cap \mathcal{P}(\omega)$ , that is, the set of all constructible sets of natural numbers. Then, if we take the second-order domain to be  $\Omega$ , we get a model of full second-order arithmetic. Here’s one way to think about why. Consider the proof that the ‘standard’ model (in which the second-order domain is the ‘real’ powerset of  $\omega$ ) actually is a model of second-order arithmetic. That proof can be formalized in ZF. But there is an inner model of ZF in which  $V = L$ , that is, in which every set is constructible, and the proof works just as well there. That is: as far as ZF is concerned,  $\mathcal{P}(\omega)$  might as well be  $\Omega$ . Moreover, the so-called “minimum  $\beta$ -model” of second-order arithmetic (Simpson, 2009, Theorem VII.5.17) satisfies the condition mentioned earlier: every set in the second-order domain of this model is (non-parametrically) definable in the model.<sup>37</sup>

Pushing forward, suppose we were to say, quite generally, that the second-order quantifiers range over the constructible powerset of the first-order domain. That is: if  $D$  is the first-order domain, then the second-order domain is  $L \cap \mathcal{P}(D)$ .<sup>38</sup> Then this would support impredicative comprehension. Moreover, the resulting logic has many of the same meta-logical properties that second-order logic standardly interpreted has, because the usual proofs of those properties go through in the inner model of ZF in which  $V = L$ : the logic is incomplete and non-compact, and Löwenheim-Skolem fails, since second-order arithmetic is categorical. Some of these arguments, however, depend upon the absoluteness of  $\omega$ ,<sup>39</sup> so it would be unwise to assume that this logic will have *all* the meta-logical properties of standard second-order logic.

<sup>36</sup> It is far from obvious, that is, if we do not allow definitions with second-order parameters. It is trivial that every such set is parametrically definable, since  $Fx$  will define  $S$  when  $S$  is the value of  $F$ .

<sup>37</sup> Thanks to John Burgess, Ali Enayat, Harvey Friedman, and Steve Simpson for responding to a question on this topic that I posted to the Foundations of Mathematics mailing list, and for subsequent discussion. Burgess, Enayat, and Friedman all sketched an argument somewhat different from Simpson’s, but that argument also makes use of an inner model where  $V = L$ .

<sup>38</sup> This assumes that  $D$  is itself constructible, which seems reasonable in the present setting. But we can make use of relative constructibility, even if  $D$  is not constructible.

<sup>39</sup> That is, the fact that we do not have different (non-isomorphic) sets playing the role of  $\omega$  in different models.

Indeed, if we really do think that the second-order quantifiers range only over the constructible powerset of the first-order domain, then we ought to endorse a number of principles as logically valid that would not normally be taken to be. For example, we should regard second-order forms of the axiom of choice as valid, since  $V = L$  implies choice.<sup>40</sup> Perhaps less encouragingly, we should regard the second-order statement corresponding to the generalized continuum hypothesis as logically valid. This would have significant consequences for any neo-Fregean reconstruction of real analysis, let alone of set theory.<sup>41</sup>

One thing this uncontroversially shows, at least assuming  $V \neq L$ , is that there are perfectly natural models of second-order logic in which the second-order domain is not the full powerset. But, while it is not unreasonable to regard constructibility, in Gödel's sense, as an idealized notion of definability, it is a *very* idealized notion. We obviously need a good deal of set-theoretic machinery even to state the thesis that the second-order domain is the constructible powerset—e.g. a robust theory of the ordinals—so this sort of view will do little to assuage any worries that second-order logic is “set theory in sheep's clothing.”

Moreover, the ‘minimal  $\beta$ -model’ of second-order arithmetic is precisely that: the *minimal* model of second-order arithmetic satisfying the conditions (i) that the model's numbers are isomorphic to the natural numbers and (ii) that the model agrees with the standard model for  $\Sigma_1^1$  formulae. It can be proven (Simpson, 2009, Theorem VII.5.17) that the minimal  $\beta$ -model is the one we get by iterating Gödel's construction up to a certain countable ordinal  $\gamma$  and then taking the second-order domain to be  $L_\gamma \cap \mathcal{P}(\omega)$ . This seems to me to imply that any reasonable model of second-order arithmetic will have to contain at least  $L_\gamma \cap \mathcal{P}(\omega)$ , i.e. that any notion of definability able to support full impredicative comprehension will have *at least* to count everything in  $L_\gamma \cap \mathcal{P}(\omega)$  as being definable. This is a tall order, even though  $L_\gamma \cap \mathcal{P}(\omega)$  is a countable set, since  $\gamma$  is, in technical parlance, *way out there*. Even to get a  $\beta$ -model of  $\Pi_1^1$  second-order arithmetic, we must go well beyond the recursive ordinals and so well beyond the ordinals for which we can have a system of notations.<sup>42</sup> I therefore find it difficult to see how any reasonable notion of a ‘possible predicate’ can do what Hale wants it to do.

## 4 Me on Schematic Logic

In “A Logic for Frege's Theorem,” I showed that what I there called “Arché logic” is sufficient for the proof of Frege's Theorem. I further argued that Arché logic has a

<sup>40</sup> Shapiro (this volume, §4.2) suggests, by contrast, that Hale might have difficulty endorsing choice: the inference from  $\forall x \exists y R(x, y)$  to  $\exists f \forall x R(x, fx)$  seems to imply, on Hale's view, that there is a *definable* function that does the trick. But the fact that choice follows from  $V = L$ , and is regarded as valid by intuitionists, counsels caution here.

<sup>41</sup> Friedman has recently shown, for example, that there are fairly simple properties that all Borel functions provably have that, if lifted to *all* real-valued functions, imply  $\neg$ CH and so imply  $V \neq L$ . (The relevant writings can be found in the archives of the Foundations of Mathematics mailing list.) For example, if  $f$  is a Borel function, then for some  $x$  and  $y$  and any  $n$ ,  $x \neq f(y + n)$  and  $y \neq f(x + n)$ . But if we assume this for all real-valued  $f$ , then  $\neg$ CH follows.

<sup>42</sup> The relevant ordinal in the case of  $\Pi_1^1$  comprehension is the supremum of the first  $\omega$  admissible ordinals, and the first admissible ordinal after  $\omega$  itself is  $\omega_1^{CK}$ , the first non-recursive ordinal (Simpson, 2009, p. 293).

significantly better claim to be ‘logic’ than does full second-order logic. Hale (2015, §5) has argued against this latter claim. More precisely, he argues that I face a dilemma: either we interpret second-order variables as ranging over the full powerset of the first-order domain, in which case Arché logic is no improvement on second-order logic as standardly interpreted; or else we interpret second-order variables as ranging over definable subsets of the first-order domain, in which case Arché logic is no improvement on second-order logic as Hale interprets it.

For our purposes, we can focus simply on the treatment of the ancestral in Arché logic.<sup>43</sup> Frege, of course, defines the ancestral in terms of second-order quantification. But we can replace Frege’s explicit definition with a pair of rules:

$$\begin{aligned} (A+) \quad & Fa \wedge \forall x(Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb \vdash Q^{*=}ab \\ (A-) \quad & Q^{*=}ab \vdash Fa \wedge \forall x(Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb. \end{aligned}$$

Here,  $F$  is a free second-order variable and, in  $(A+)$ , it is required not to be free in any assumption on which the premise of the rule depends. The key technical point is that we can state these rules, and reason with them, in a system that does not even permit second-order quantification. We allow for *free* second-order variables, but not for bound ones.

As I have emphasized, whereas Hale and Wright seem primarily to be concerned with the ontology of second-order logic, my concern is primarily with its epistemology. Thus I write:

[T]he crucial issue for the Neo-logician is epistemological. The proof of Frege’s Theorem makes heavy use of the ancestral and of inferences of the sort described by  $(A+)$  and  $(A-)$ . A Neo-logician must therefore show that she is entitled both to a grasp of the concept of the ancestral and to an appreciation of the validity of  $(A+)$  and  $(A-)$ , and this entitlement must be epistemologically innocent in the sense that it does not itself import epistemological presuppositions that undermine the Neo-logician project. (Heck, 2011b, p. 286, notation altered)

My strategy is to follow Frege’s example by recasting the epistemological issue as a semantic one. What I try to show is that one can understand, and appreciate the validity of, the mode of definition that Arché logic exploits without having *any* positive conception of what the range of the second-order variables might be and so, in particular, without regarding those variables as ranging over the full powerset.

There are two pressing questions about this proposal. The first is how formulae such as the one that occurs as the premise of  $(A+)$ —formulae containing free second-order variables—are to be understood. The second is why we should regard ourselves as entitled to define the ancestral by stipulating the rules  $(A+)$  and  $(A-)$ . In fact, however, if we have just those two rules, the system is very weak: too weak for the proof of Frege’s Theorem. To get any strength, we need to liberalize  $(A-)$  as follows:<sup>44</sup>

<sup>43</sup> Full Arché logic simply generalizes this treatment of the ancestral to other notions defined in terms of  $\Pi_1^1$  formulae.

<sup>44</sup> In the earlier paper, I speak instead of a rule of substitution. That is formally equivalent to liberalizing the elimination rule, but here it is much simpler to speak of generalized rules such as  $(A-^I)$ .

$$(A^{-I}) \quad Q^{*=} ab \vdash \phi(a) \wedge \forall x (\phi(x) \wedge Qxy \rightarrow \phi(y)) \rightarrow \phi(b).$$

The difference is that, in  $(A^{-I})$ ,  $\phi(x)$  is an arbitrary *formula* and not just a variable. Thus,  $(A^{-I})$  is impredicative: it would, for example, allow us to take  $\phi(x)$  to be  $Q^{*=} ax \wedge Fx$ , as we need to do in the proof of induction. So what we need to show is that we are entitled not just to the weaker rule  $(A-)$  but also to the stronger one  $(A^{-I})$ .

Concerning the first question: it is implicit already in our capacity for *first-order* reasoning that we should be able to reason ‘schematically’ about an arbitrary Concept. Suppose, for example, that I hear some people arguing about whether Alex is meretricious. If I have no idea what that means, that might prevent me from fully evaluating their reasoning. But it will not prevent me from evaluating their *logical* reasoning, since first-order reasoning is formal. To a significant extent, I could even engage in such reasoning myself, using the term “meretricious” without understanding it. The ability to reason schematically, using free second-order variables, is just a short step beyond this: one simply lets  $F$  (or some nonsense term) go proxy for some meaningful predicate or other. What I am calling ‘schematic reasoning’ thus generalizes ordinary first-order reasoning by allowing us to abstract from the specific Concepts that are under discussion. The conclusions we thereby reach are ones that hold for any Concept: given a predicate that expresses that Concept, we can simply replicate the reasoning that led to the schematic conclusion.

One might respond that this is really substitutional reasoning, but there is nothing in the nature of schematic reasoning that requires whatever Concept is involved to be expressed by some predicate of our current language; moreover, anything in the reasoning that did depend upon such an assumption would be invalid. We are just thinking, in an abstract way, about some Concept or other—one that is specific but unspecified. As for what a Concept is, we can think of it simply as the sort of thing an object may or may not ‘have’ or, to put it in the formal mode, as the sort of thing that a predicate may express. Note that this agrees, to a significant extent, with Hale’s conception. The difference is that I see no need to require that every Concept can be expressed by a ‘possible predicate’, not if that restriction would have any bite. Indeed, as I have already said, nothing in ordinary first-order reasoning, or its schematic generalization, reflects such a restriction, and anything that did would be invalid.<sup>45</sup>

How is such reasoning different from second-order reasoning? We shall return to this question below, but the short answer is that it is conceptually far simpler. As anyone who has ever taught introductory logic well knows, what makes quantificational reasoning complex is not the use of variables. That is usually familiar to students from basic algebra. What makes quantification both difficult and powerful is the phenomenon of *scope* and, more significantly, the possibility that one quantifier may *depend upon* another. The lack of second-order quantifiers in schematic logic makes such dependencies impossible to express.

<sup>45</sup> To clarify something (see Hale, 2015, p. 18): I do *not* claim that formulae of the form  $\dots F \dots$  are always to be understood in terms of the so-called “closure interpretation,” i.e. as equivalent to  $\forall F(\dots F \dots)$ . Indeed, I make it quite clear that they are not (Heck, 2011b, pp. 280–1, esp. fn. 29). Rather, such reasoning should be understood as involving an arbitrary Concept—specific but unspecified.

The second question was what underlies our appreciation of the validity of the rules of inference governing the ancestral. What I would like to be able to argue is that these rules may legitimately be treated as *defining* the ancestral, so that the rules are self-justifying (Heck, 2011b, p. 286). Such a view rests upon what one might call “local inferentialism.” It’s ‘local’ because it insists that this particular kind of definition is a legitimate way of introducing a new predicate and bestowing meaning upon it. That claim is not supposed to fall out of a general view to the effect that content is always bestowed by inferential role and that we can make our words mean what we like so we can stipulate any rules we please.

But the truth is that I do not argue for this sort of thesis in the earlier paper, and I am not going to do so here, either. My goal, now as then, is more limited: it is to show that our entitlement to these rules, and in particular to the impredicative rule ( $A-I$ ), does not depend upon any presuppositions connected with the range of the second-order quantifiers. It emerges, rather, from the idea that schematic reasoning is unrestrictedly general. To put it differently, the idea is to argue that, if we can earn a right to the restricted rule

$$(A-) \quad Q^{*=}ab \vdash Fa \wedge \forall x(Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb$$

then we can earn a further right to the unrestricted rule

$$(A-I) \quad Q^{*=}ab \vdash \phi(a) \wedge \forall x(\phi(x) \wedge Qxy \rightarrow \phi(y)) \rightarrow \phi(b)$$

without having to invoke the assumption that the second-order variables range over the full powerset of the first-order domain. This is because schematic reasoning concerns *any Concept whatsoever*: nothing in the nature of that reasoning depends upon any assumption about how a given Concept can be specified, or even whether it can be specified at all. If so, then it is just irrelevant what conceptual resources are deployed in the predicate  $\phi(\xi)$ : if we know that  $Fa \wedge \forall x(Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb$  holds no matter what Concept  $F$  might be, then it holds in the particular case of the Concept expressed by  $\phi(\xi)$ .

That, then, is the view developed in “A Logic for Frege’s Theorem.”

Hale begins his criticism of this view by discussing my response to the objection that our understanding of the introduction rule for the ancestral involves a conception of the full power-set of the domain. How else, it might be asked, are we to understand

$$(**) \quad Fa \wedge \forall x \forall y (Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb$$

as it occurs in the premise of the rule ( $A+$ ), except as involving a tacit initial second-order quantifier? [1] Does it not say, explicit quantifier or no, that all concepts  $F$  that are thus-and-so are so-and-thus? [2] Doesn’t understanding that claim therefore require the disputed conception of the full power-set?

(Heck, 2011b, p. 293, example and notation changed, indices added)

Hale (2015, p. 18) accuses me of running questions (1) and (2) together when I respond:

No, it does not. A better reading would be: A concept that is thus-and-so is so-and-thus. What understanding this claim requires is not a capacity to conceive of *all* concepts but simply the capacity to conceive of *a* concept: to conceive of an arbitrary concept, if you like.

(Heck, 2011b, p. 293, emphasis original).

But I do not think I was running the two questions together. These remarks were intended only as a response to question (1): how (\*\*) is to be read. Hale may have been misled by the fact that I never do answer question (2). That is because my response to question (1) was supposed to make question (2) moot, since question (2) presupposes that the answer to question (1) was affirmative.

The objection we are discussing is motivated by the observation that, when (\*\*) occurs as the premise of (A+), the variable  $F$  is required not to be free in any assumption on which (\*\*) depends. So, just as in the case of Universal Generalization, we will be able to infer  $Q^{*}ab$  from (\*\*) only if the universal closure of (\*\*) is true. But that does not invalidate the reading according to which (\*\*) concerns a specific but unspecified Concept. Requiring that  $F$  not be free in any assumption on which (\*\*) depends is simply a way of registering, at the formal level, the intuitive requirement that  $F$  has to be ‘arbitrary,’ i.e. that, if the intended application of (A+) is to be legitimate, we cannot have made any specific assumptions about  $F$  in reaching (\*\*).<sup>46</sup>

Now, to be sure, if we want to give a proper semantic account of such languages, we will need to make use of the notion of assigning a value to such variables, and we’ll need to quantify over such values, too. In particular, in this case, we’ll want to say that the inference is legitimate only if the premise is true for every assignment to  $F$ . But the resources we need here are ones we already need to give a semantics for *first-order* logic, at least when it comes to defining validity and implication:<sup>47</sup> we need the idea of a predicate-letter’s having an interpretation, and we need to quantify over the possible values. We need nothing more to define validity and implication for the language of schematic logic. And, as Wright (2007, pp. 162–3) emphasizes in his discussion, we need to distinguish what is required to give a *semantics* for such statements from what is involved in *understanding* them. I am claiming that we can *understand* such statements in terms of the notion of what is true of an arbitrary—specific but unspecified—Concept and that we do not need to have even an inkling of what it might mean to say that something is true of *all* Concepts.

My imaginary interlocutor, in my earlier discussion, was not satisfied with this sort of reply. They continued to press their objection by insisting that, “if we do not have a definite conception of the full power-set of the domain . . . then the meanings of the predicates we introduce by schematic definition will be radically underdetermined” (Heck, 2011b, p. 294). Making the distinction on which I’ve just insisted does not change the fact that  $Q^{*}ab$  is true just in case  $Fa \wedge \forall x(Fx \wedge Qxy \rightarrow Fy) \rightarrow Fb$  is true for every assignment to  $F$ . So the truth-conditions of  $Q^{*}ab$  depend upon what exactly the second-order domain comprises, and there will be plenty of choices, hence the indeterminacy. My response to this version of the objection is to insist, once again, on the unrestricted generality of our understanding of schematic reasoning: there is only one possible choice, since the second-order variables must include in their range everything that they possibly could include.

<sup>46</sup> Wright (2007, pp. 156–7) seems to envisage a treatment on which  $F$  might actually be an ordinary atomic predicate, but we impose a restriction to the effect that the derivation of the premise does not depend upon which predicate it is. That would do here as well.

<sup>47</sup> Antonelli (2013) has argued, in fact, that an appropriately general semantics for *first-order* quantifiers depends upon a notion of a *second-order* domain.



Taking up my interlocutor's cause, Hale (2015, pp. 22–3) objects that the unrestricted generality of schematic reasoning, even if accepted, does not by itself imply that the second-order variables range over the full powerset. But it was never intended to do so. To the contrary, what I am trying to argue is that we do *not* need to conceive of the second-order domain as being the full powerset in order to avoid the threatened indeterminacy. More generally, I am arguing that we can avoid the indeterminacy even if we have *no positive conception whatsoever* of what the second-order domain comprises.

There can be no disputing that axiomatic higher-order logics will always have models with different second-order domains, even once we fix the first-order domain.<sup>48</sup> The question is whether this agreed fact implies that the axioms and rules we accept do not fix a 'privileged' model in terms of which the actual truth-conditions are determined. If they do not, the worry goes, then the only way that definite truth-conditions could be fixed is if we had some positive conception of what the second-order domain was to be—one that, of necessity, was not wholly grounded in our understanding of the axioms and rules. To many, that has seemed bad. But the mere fact that there are many models of the axioms and rules we accept simply does not imply that our understanding of those axioms and rules does not fix a privileged model.

Assume for the moment that the privileged model is the one whose second-order domain contains every subset of the first-order domain. Then the issue takes the following specific shape. It is a mathematical fact that the axioms and rules of Arché logic do not require the second-order domain of all of its models to be the full powerset. My interlocutor therefore wants to conclude that the axioms and rules of Arché logic do not require the second-order domain to be the privileged domain. I mean to be challenging that inference. The way we understand the axioms and rules of Arché logic might require the second-order domain to be a certain privileged domain *without* requiring it to be the full powerset, *even if the privileged domain is the full powerset*.

The unrestricted character of our understanding of schematic reasoning forces the privileged model to be the *maximal* model: one in which the domain includes everything it could possibly include. Maybe that will ultimately imply that the second-order domain is the full powerset. Even if so, however, there is no reason that further fact needs to be implicit in our understanding of schematic reasoning. One might instead regard questions about the nature of the privileged model as significant philosophical, mathematical, and semantic questions that simply are not resolved by our ordinary understanding of the axioms and rules of Arché logic, even though that understanding does require the privileged domain to be maximal.

What certainly *does* matter is that our understanding of the axioms and rules should support the impredicative form ( $A^{-I}$ ) of the elimination rule for the ancestral. As we have seen, it is often assumed that the only 'natural' assumption that will do the necessary work is that the second-order domain comprises the full powerset. In response, I am arguing not only that this assumption is not required but that no alternative positive conception of the domain is required either. It is enough that our understanding of the rules governing the ancestral involves the assumption that the

<sup>48</sup> Assuming, of course, that the first-order domain is not finite, and meaning by 'model' anything that validates all the axioms and rules.

application of these rules is unrestricted: that, in particular, we accept no restriction, in applications of  $(A-I)$ , to predicates definable in terms of ‘more basic’ notions. In effect, then, what I need is only the much less controversial half of Hale’s view: that any definable concept will be in the domain. I have no need for, and see no reason to commit myself to, the other half of his view: that *only* definable concepts are in the domain.

Hale does want to make the stronger claim, of course, because he is particularly concerned with worries about ontological extravagance. But I have no such worries, and I doubt that they are at the root of Quine’s objections, either. I think Quine’s central worry is that our *understanding* of second-order logic is parasitic on our *understanding* of set theory: even if second-order variables do not range over ‘classes,’ then our appreciation of what they do range over depends upon our understanding that, for every subclass of the first-order domain, there is some element of the second-order domain that corresponds to it. That is what I am trying to fight off. And Hale and I are largely in agreement how to do it: our understanding of second-order variables only requires us “to know what general condition something must satisfy, if it is to be a possible value” (Hale, 2013, p. 154). And we are largely agreed, too, what that condition is: being the sort of thing for which a predicate can stand. Where we differ, again, is that Hale wishes to impose an additional condition: that a Concept exists only if there could be some predicate that expressed it. I, by contrast, find that claim unhelpful because unclear, and I simply see no reason to deny, with Hale, that every subset of the first-order domain corresponds to something in the range of the second-order variables.

## 5 Schematic Logic versus Second-Order Logic

The obvious question, at this point, is whether the sorts of resources I’ve allowed myself can be deployed to justify full second-order logic. In particular, one might wonder why we cannot just follow Wright by stipulating the usual introduction and elimination rules for the second-order quantifiers and then appeal to the unrestricted character of our understanding of schematic reasoning to justify impredicative comprehension.

The answer is that there is an important disanalogy. In order to get the full power of second-order logic, we need to allow *parameters* in the introduction and elimination rules. That is, we need to allow inferences of the form:

$$\dots F \dots G \dots \vdash \forall F(\dots F \dots G \dots).$$

Without these, one will not be able to handle cases in which one quantifier is embedded inside another. In the case of Arché logic, by contrast, such parameters are not permitted. We could generalize the method by which the ancestral was introduced, if we wished, to allow such definitions as:

$$\begin{aligned} \dots F \dots G \dots x \dots \vdash \Phi_y(Gy, x) \\ \Phi_y(Gy, x) \vdash \dots F \dots G \dots x \dots \end{aligned}$$

If we do permit this sort of definition, then we get a system with the power of full second-order logic (Heck, 2011b, pp. 292–3). But now this really is very different:  $G$  is here being treated as an *argument* of the newly introduced predicate  $\Phi$ , which is

for that reason not a first-order predicate. The semantics of this sort of language thus requires a notion of a higher-order Concept, specifically, in this case, a Concept of mixed level under which Concept-object pairs do or do not fall. Arché logic has no need for any such notion.

In the case of second-order logic as usually formulated, the difference is felt elsewhere: in how the values of the second-order variables are handled. In a semantic theory for a second-order language, we need to include values for second-order variables in the sequences with respect to which we define satisfaction. Their values cannot be fixed once and for all but may change as we break a formula down into its parts. In the case of Arché logic, by contrast, we need only define truth with respect to some *fixed* assignment of values to the second-order variables: in effect, we can treat second-order variables the way predicate-letters are treated in first-order logic. And this difference, in turn, has consequences concerning the expressive power of the two languages.

It is well known that there are many important arithmetical facts that monadic first-order logic will not allow us to express, for example, that there is no greatest number. This is, in part, due to the restriction to monadic predicates:  $\forall x\exists y(x < y)$  essentially involves the relational predicate  $\xi < \zeta$ . But what is really crucial is not so much the lack of relational predicates as the inability of monadic first-order logic to capture dependency relations between quantifiers. We can write such things as  $\forall x(Fx \rightarrow \exists y(Gy \wedge Hx))$ , with the existential quantifier in the scope of the universal one. But this is logically equivalent to  $\forall x(Fx \rightarrow Hx) \wedge [\exists x(Fx) \rightarrow \exists y(Gy)]$ . There is thus no real dependency: the value of  $y$  does not depend upon that of  $x$ . That is what really matters. That is what scope is really about. It would do us no good to allow relational predicates if, say, we permitted only  $x$  to occur as a bound variable. That would prevent one quantifier from occurring in the scope of another, so we'd still be unable to express such dependencies and, *a fortiori*, be unable to capture the difference between  $\forall x\exists y(x < y)$  and  $\exists y\forall x(x < y)$ .

Arché logic has exactly the same sort of limitation. There is obviously a sense in which schematic definitions, like that of  $Q^* = ab$ , allow us to get the effect of  $\forall F(\dots F \dots)$ . But they will *not* allow us to get the effect of  $\forall F(\dots F \dots G \dots)$ , because second-order parameters are not permitted. As a result, there is no way, in Arché logic, to express something of the form  $\forall F\exists G(\dots F \dots G \dots)$ , not if the value of  $G$  actually depends upon the value of  $F$ . For this reason, Arché logic is not just logically weaker than full second-order logic but *expressively* weaker: there are things we *cannot say* in the language of Arché logic that we can say in the language of second-order logic. And that is the most important sense, it seems to me, in which Arché logic is not just second-order logic in unfamiliar clothing. It is also why, contra Hale (2015, §5.3), I do not think it a merely terminological question whether the language of Arché logic is a second-order language.

## 6 Closing

It would be premature to draw any firm conclusions from our discussion. As we saw in Section 3, it is not easy to find a notion of definability that will do the sort of work Hale needs it to do. It turns out that there is such a notion, one we can

characterize in terms of Gödel's notion of a constructible set. But that notion is very idealized and intimately bound to broadly set-theoretic notions: it is itself defined in terms of iterating a certain process through the ordinals, and not just a few of them, either. As a result, it is doubtful that taking the second-order domain to be the constructible powerset would have any epistemological advantage over taking it to be the full powerset.

My own response to this situation has been to scale back my ambitions. We do not actually need full second-order logic to count as 'logic' for Frege's Theorem to have the sort of epistemological significance some of us have hoped it might. In particular, we do not need full comprehension: it is enough if  $\Pi_1^1$  comprehension counts as logically valid. The problem is that it is not entirely obvious how to make philosophy out of this technical fact. One way might be to try to articulate a notion of definability that, even if it did not support full comprehension, would at least support  $\Pi_1^1$  comprehension. I don't necessarily say that isn't possible.<sup>49</sup> But it is also possible to approach the problem less semantically and more syntactically: don't worry so much about what the domain is; try instead to make it plausible that certain axioms and rules are epistemologically innocent. That is the strategy I have been pursuing. Whether it is the right one is, of course, a matter for continued debate.<sup>50</sup>

## References

- Antonelli, A. (2013). "On the general interpretation of first-order quantifiers," *Review of Symbolic Logic* 6: 637–58.
- Barwise, J. and Cooper, R. (1981). "Generalized quantifiers and natural language," *Linguistics and Philosophy* 4: 159–219.
- Benacerraf, P. (1995). "Frege: The last logicist," in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics*. Cambridge, MA, Harvard University Press, 41–67.
- Boolos, G. (1998). "Is Hume's principle analytic?" in R. Jeffrey (ed.), *Logic, Logic, and Logic*. Cambridge, MA, Harvard University Press, 301–14.
- Boolos, G. and Heck, R. G. (2011). "Die Grundlagen der Arithmetik §§82–83," in Heck 2011a, 69–87.
- Büchi, J. R. (1962). "On a decision method in restricted second order arithmetic," in E. Nagel et al. (eds), *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress*. Stanford, CA, Stanford University Press, 1–11.
- Burgess, J. P. (2005). *Fixing Frege*. Princeton, NJ, Princeton University Press.
- Cook, R. (2014). "Possible predicates and actual properties," *Synthese*, doi: 10.1007/s11229-014-0592-1.

<sup>49</sup> Getting a model of  $\Pi_1^1$  second-order arithmetic requires a lot less than getting a model of full second-order arithmetic: we only need the second-order domain to be closed under any finite number of iterations of the hyperjump, and under relative recursive definability (Simpson, 2009, Corollary VII.1.9).

<sup>50</sup> Thanks to Stewart Shapiro and James Walsh for helping me think through some of these issues, and to the editors for their comments upon a draft of the paper.

It is the peculiar tradition of our tribe to express our respect for other members by highlighting our disagreements with them. So, in case it is not clear, let me just say explicitly how much I admire Bob Hale's work. I learned a lot from him over the years—both in conversation and from his written work—and greatly enjoyed the time we were able to spend together. Bob's enthusiastic support for me and my work, early in my career, was particularly important to me. So I am honored to be able to contribute to this volume and thank Ivette and Jessica for the invitation.

- Dedekind, R. (1902). "The nature and meaning of numbers," trans. W. W. Beman, in *Essays on the Theory of Numbers*. Chicago, Open Court, 31–115.
- Dummett, M. (1981). *Frege: Philosophy of Language*, 2nd edition. Cambridge, MA, Harvard University Press.
- Eklund, M. (1996). "On how logic became first-order," *Nordic Journal of Philosophical Logic* 1: 147–67.
- Enderton, H. (2001). *A Mathematical Introduction to Logic*, 2nd edition. London, Academic Press.
- Evans, G. (1985). "Semantic structure and logical form," in *Collected Papers*. Oxford, Clarendon Press, 49–75.
- Feferman, S. (1999). "Logic, logics, and logicism," *Notre Dame Journal of Formal Logic* 40: 31–54.
- Ferreirós, J. (2001). "The road to modern logic: An interpretation," *Bulletin of Symbolic Logic* 7: 441–84.
- Frege, G. (1980). *The Foundations of Arithmetic*, 2nd revised edition, trans. J. L. Austin. Evanston, IL, Northwestern University Press.
- Frege, G. (2013). *The Basic Laws of Arithmetic*, trans. P. A. Ebert and M. Rossberg. Oxford, Oxford University Press.
- Gödel, K. (1938). "The consistency of the axiom of choice and of the generalized continuum-hypothesis," *Proceedings of the National Academy of Sciences of the United States of America* 24: 556–7.
- Hale, B. (2013). "Properties and the interpretation of second-order logic," *Philosophia Mathematica* 21: 133–56.
- Hale, B. (2014). "Critical notice of Richard Heck's *Frege's Theorem*," *Mind* 123: 437–56.
- Hale, B. (2015). "Second-order logic: Properties, semantics, and existential commitments," *Synthese*, doi: 10.1007/s11229-015-0764-7.
- Heck, R. G. (2011a). *Frege's Theorem*. Oxford, Clarendon Press.
- Heck, R. G. (2011b). "A logic for Frege's Theorem," in Heck 2011a, 267–96.
- Heck, R. G. (2011c). "Ramified Frege arithmetic," *Journal of Philosophical Logic* 40: 715–35.
- Heck, R. G. (2011d). "Syntactic reductionism," in Heck 2011a, 180–99.
- Heck, R. G. (2012). *Reading Frege's Grundgesetze*. Oxford, Clarendon Press.
- Heck, R. G. (2014). "Frege arithmetic and 'everyday mathematics,'" *Philosophia Mathematica* 22: 279–307.
- Heck, R. G. (2016a). "Formal arithmetic before *Grundgesetze*," in P. Ebert and M. Rossberg (eds), *A Companion to Grundgesetze*. Oxford, Oxford University Press.
- Heck, R. G. (2016b). "Is Frege's definition of the ancestral adequate?" *Philosophia Mathematica* 24: 91–116.
- Kerry, B. (1887). "Über Anschauung und ihre psychische Verarbeitung, vierter Artikel," *Vierteljahrsschrift für wissenschaftliche Philosophie* 11: 249–307.
- Koellner, P. (2010). "Strong logics of first and second order," *Bulletin of Symbolic Logic* 16: 1–36.
- Kripke, S. (1975). "Outline of a theory of truth," *Journal of Philosophy* 72: 690–716.
- Linnebo, Ø. (2004). "Predicative fragments of Frege arithmetic," *Bulletin of Symbolic Logic* 10: 153–74.
- McLarty, C. (2013). "Interpreting set theory in higher order arithmetic," <http://arxiv.org/abs/1207.6357>.
- Moore, G. H. (1988). "The emergence of first-order logic," *Minnesota Studies in the Philosophy of Science* 11: 95–135.
- Mostowski, A. (1957). "On a generalization of quantifiers," *Fundamenta Mathematicae* 44: 12–36.

- Parsons, C. (1974). "Sets and classes," *Noûs* 8: 1–12.
- Parsons, C. (1981). "Frege's theory of number," in *Mathematics in Philosophy*. Ithaca, NY, Cornell University Press, 150–75.
- Quine, W. V. O. (1970). *Philosophy of Logic*. Englewood Cliffs, NJ, Prentice Hall.
- Quine, W. V. O. (1986). *Philosophy of Logic*, 2nd edition. Cambridge, MA, Harvard University Press.
- Rayo, A. and Uzquiano, G. (1999). "Toward a theory of second-order consequence," *Notre Dame Journal of Formal Logic* 40: 315–25.
- Rayo, A. and Yablo, S. (2001). "Nominalism through de-nominalization," *Noûs* 35: 74–92.
- Russell, B. (1908). "Mathematical logic as based on the theory of types," *American Journal of Mathematics* 30: 222–62.
- Shapiro, S. (1991). *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford, Oxford University Press.
- Simpson, S. (2009). *Subsystems of Second Order Arithmetic*, 2nd edition. Cambridge, Cambridge University Press.
- Tarski, A. (1958a). "The concept of truth in formalized languages," in Tarski 1958b, 152–278.
- Tarski, A. (1958b). *Logic, Semantics, and Metamathematics*, ed. J. Corcoran. Indianapolis, IN, Hackett.
- Tarski, A. (1958c). "On definable sets of real numbers," in Tarski 1958b, 110–42.
- Visser, A. (2009). "The predicative Frege hierarchy," *Annals of Pure and Applied Logic* 160: 129–53.
- Westerståhl, D. (2015). "Generalized quantifiers," *Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/sum2015/entries/generalized-quantifiers/>.
- Wright, C. (1983). *Frege's Conception of Numbers as Objects*. Aberdeen, Aberdeen University Press.
- Wright, C. (2007). "On quantifying into predicate position," in M. Leng (ed.), *Mathematical Knowledge*. Oxford, Oxford University Press, 150–74.