Logicism Reconsidered*

Agustín Rayo
ar29@st-andrews.ac.uk

November 28, 2002

This paper is divided into four sections. The first two identify different logicist theses, and show that their truth-values can be conclusively established on minimal assumptions. Section 3 sets forth a notion of 'content-recarving' as a possible constraint on logicist theses. Section 4—which is largely independent from the rest of the paper—is a discussion of 'Neo-Logicism'.

1 Logicism

1.1 What is Logicism?

Briefly, logicism is the view that mathematics is a part of logic. But this formulation is imprecise because it fails to distinguish between the following three claims:

1. *Language-Logicism*

   The language of mathematics consists of purely logical expressions.

2. *Consequence-Logicism*

   There is a consistent, recursive set of axioms of which every mathematical truth is a purely logical consequence.

3. *Truth-Logicism*

   Mathematical truths are true as a matter of pure logic.

In fact, this is still not what we want. Consider *Language-Logicism* as an example. When taken at face value, standard mathematical languages do not consist of purely logical expressions. (For instance, when taken at face value, the standard language of arithmetic contains the individual constant '0', which is not a purely logical expression.) So if *Language-Logicism* is to have any plausibility, it should be read as the claim that mathematical sentences can be *paraphrased* in such a way that they contain no non-logical vocabulary. Similarly for *Consequence-Logicism* and *Truth-Logicism*: if they are to have any

*I am grateful to Matti Eklund, Stewart Shapiro and Gabriel Uzquiano for many helpful comments.

plausibility, they should be read, respectively, as the claim that mathematical truths can be *paraphrased* in such a way that they are a purely logical consequence of the relevant set of axioms, and as the claim that mathematical truths can be *paraphrased* in such a way that their truth is a matter of pure logic. Let us therefore distinguish between the following logicist theses (relative to a mathematical language $\mathcal{L}$ with an intended model $\mathcal{M}$):

> *Notation:* A *paraphrase-function* $*$ is a function such that, for any sentence ϕ in the domain of $*$, ϕ^* is a paraphrase of ϕ.

1. *Language-Logicism*

   There is a paraphrase-function $*$ such that, for any sentence ϕ of $\mathcal{L}$, ϕ^* contains no non-logical vocabulary.

2. *Consequence-Logicism (Semantic Version)*

   There is a paraphrase-function $*$ and a consistent, recursive set of sentences $\mathcal{A}$ such that, for any sentence ϕ of $\mathcal{L}$ which is true (false) according to $\mathcal{M}$, ϕ^* (the negation of ϕ^*) is a semantic consequence of $\mathcal{A}$.

3. *Consequence-Logicism (Syntactic Version)*

   Like the semantic version, except that 'semantic consequence of $\mathcal{A}$' is replaced by 'derivable from $\mathcal{A}$ on the basis of purely logical axioms and rules of inference'.

4. *Truth-Logicism (Semantic Version)*

   There is a paraphrase-function $*$ such that, for any sentence ϕ of $\mathcal{L}$ which is true (false) according to $\mathcal{M}$, ϕ^* (the negation of ϕ^*) is a logical truth.

5. *Truth-Logicism (Syntactic Version)*

   Like the semantic version, except that 'logical truth' is replaced by 'derivable from the empty set on the basis of purely logical axioms and rules of inference'.

In stating the five logicist theses, we have left open two important issues. First, there is the question of what it takes for one sentence to be a *paraphrase* of another. Is synonymy required, or could we settle for less? Second, there is the question of what should be counted as *logic*. Which expressions count as logical? Which axioms and rules of inference are logical? Which sentences are logical truths? These are difficult questions, which we cannot hope to address here. In section 2 we will evaluate the logicist theses on the basis of different assumptions about logic and paraphrase, but refrain from assessing the assumptions themselves.

1.2 Why Bother?

The five logicist theses are connected to important philosophical problems. Here are some examples:

1. If *Language-Logicism* is true, it might be used as part of a nominalist account of mathematics.[1]

2. (a) If the semantic version of *Consequence-Logicism* is true, then every mathematical sentence has a determinate truth-value, provided only that the relevant axioms are determinately true and that semantic consequence preserves determinate truth. Such a result is important for certain brands of structuralism, which face the challenge of explaining how it is that finite beings like ourselves can succeed in picking-out the standard mathematical structures. (See 'Structuralism' and 'Structuralism Reconsidered', in this volume.)

   (b) If, in addition, the syntactic version of *Consequence-Logicism* is true, then one can come to know (at least in principle) any mathematical truth by deriving it from an appropriate set of axioms, provided only that the axioms are known to be true and that derivations preserve knowledge.

3. (a) If the semantic version of *Truth-Logicism* holds, then every mathematical truth is a logical truth.

   (b) If, in addition, the syntactic version of *Truth-Logicism* holds, then we get a very impressive result: one can come to know (at least in principle) any mathematical truth by carrying out a purely logical derivation, provided only that logical derivations preserve knowledge.

1.3 Some Brief Historical Remarks

An important focus of traditional logicist projects was *Truth-Logicism*. Famously, Frege's *Foundations of Arithmetic* and *Basic Laws of Arithmetic* attempted to establish a version of *Truth-Logicism* according to which mathematical truths can be proved merely on the basis of 'general logical laws and definitions'. The precise motivation behind Frege's defense of *Truth-Logicism* is a matter of scholarly debate.[2]

Frege's project collapsed with the discovery that one of its Basic Laws leads to contradiction. In the latter part of the twentieth century, however, the project was revitalized in a somewhat different form by Crispin Wright and Bob Hale, largely as an attempt to secure a version of the epistemological result in 3(b). (For more on Neo-Fregeanism see section 4 of the present text and 'Logicism in the 21st Century', in this volume.)

The logical empiricists of the first half of the twentieth century were also interested in *Truth-Logicism*. Their program was based on the doctrine that any

[1] For more on nominalistic accounts of mathematics, see Burgess and Rosen (1997).
[2] See, for instance, Benacerraf (1981).

meaningful statement concerns either a 'matter of fact' (and, if true, is knowable only *a posteriori*), or a 'relation of ideas' (and, if true, is 'true in virtue of meaning' and consequently knowable *a priori*).[3] In order to resist the idea that mathematical statements concern 'matters of fact', it was therefore crucial to logical empiricism that every mathematical truth be shown to be 'true in virtue of meaning' and consequently knowable *a priori*. But if every mathematical truth can be known on the basis of a purely logical derivation, as in 3(b), then mathematical truths are indeed 'true in virtue of meaning' and knowable *a priori*—assuming that logical truths are 'true in virtue of meaning' and that purely logical derivations preserve *a priori* knowledge, as logical empiricists believed.

It is not clear, on the other hand, whether *Truth-Logicism* was Russell's aim in *Principia Mathematica*. It might well have consisted of *Language-Logicism* together with *Consequence-Logicism*.[4] (For more on the logicism of Frege and Russell, see 'Logicism', in this volume.)

2 An Assessment of Logicism

2.1 Logic and Paraphrase

In stating our five logicist theses in section 1.1, we left open the question of what should be counted as logic. Two possible answers are that logic is first-order logic, and that logic is higher-order logic. More precisely,

> *The First-Order View*
> A sentence is a logical truth just in case it can be paraphrased as a first-order sentence which is true in every model of the standard first-order semantics.[5] An axiom is logical just in case it is a logical truth. A rule of inference is logical just in case it preserves logical truth. Logical expressions are those that can be paraphrased as an expression of pure first-order logic.

> *The Higher-Order View*
> Like the first-order view, except that 'first-order' is everywhere replaced by 'higher-order'.[6] (For more on higher-order languages, see

[3]See, for instance, the preface to the first edition of Ayer (1946).

[4]See, for instance, Boolos (1994)

[5]Similarly, a sentence ϕ is a semantic consequence of a set of sentences Γ just in case there is a first-order sentence ϕ' and a set of first-order sentences Γ' such that (a) ϕ' is a paraphrase of ϕ, (b) every sentence in Γ' is a paraphrase of some sentence in Γ, and (c) ϕ' is true in every model of the standard first-order semantics which makes every sentence in Γ' true. In order not to count '$\exists x(x = x)$' as a logical truth, we must avoid the simplifying assumption that the domain of a model is always non-empty.

[6]Unless a suitable Reflection Principle holds, the standard semantics for higher-order languages is inadequate for languages whose intended domain contains too many objects to form a set (such as the language of set-theory). To avoid the problem, one can use a higher-order semantics, such as the one developed in Rayo and Uzquiano (1999).

'Higher-order Logic' and 'Higher-order Logic is not Logic', in this volume.)

In order to read the logicist theses in accordance with the first-order view, all we need to do is insist that the relevant paraphrase-functions deliver first-order sentences as outputs, and understand logical notions in accordance with the standard semantics for first-order languages. Similarly for the higher-order view.

These are certainly not the only ways of answering the question of what should be counted as logic—one might think, for instance, that an infinitely long sentence can be logically true.[7] But they are the only ones we will consider here.

A second issue which we left open is the question of what it takes, in the context of logicism, for one sentence to count as a paraphrase of another. This is a difficult problem. But, whatever the constraints on paraphrases turn out to be, it is reasonable to expect that any legitimate paraphrase-function will be recursive and will preserve truth-values. And we will see there is a lot that can be learned about the logicist theses on the basis of these rather minimal constraints.

Formally, let us say that a paraphrase-function $*$ is *minimally adequate* for a mathematical language $\mathcal{L}$ with an intended model $\mathcal{M}$ just in case the following conditions are satisfied:

1. Every sentence of $\mathcal{L}$ is in the domain of $*$.

2. The restriction of $*$ to sentences of $\mathcal{L}$ is recursive.

3. There is a model $\mathcal{S}$ (intuitively, the intended model of the paraphrases) such that, for any sentence ϕ of $\mathcal{L}$, $\models_{\mathcal{M}} \phi$ if and only if $\models_{\mathcal{S}} \phi^*$.

2.2 The First-order View

When the logicist theses are read in accordance with the first-order view, the following relations obtain: (see figure 1)

[FOV-1] The syntactic and semantic versions of *Truth-Logicism* and *Consequence-Logicism* are equivalent, since first-order logic is sound and complete according to the first-order view.

[FOV-2] *Truth-Logicism* implies *Consequence-Logicism*, since *Truth-Logicism* is the special case of *Consequence-Logicism* when $\mathcal{A}$ is empty.

[FOV-3] When paraphrase-functions are assumed to be minimally adequate, *Language-Logicism* implies *Consequence-Logicism*.[8]

[7] Yablo (2002) provides an interesting paraphrase-function from arithmetic and set-theory to an infinitary version of first-order logic.

[8] This follows from the fact that any consistent first-order theory T containing only sentences with no non-logical vocabulary has a recursive axiomatization: if T has a finite model,

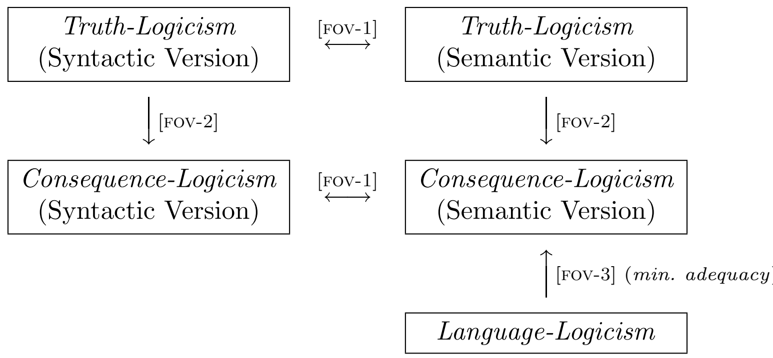

Figure 1: Implications amongst the five logicist theses, when read in accordance with the first-order view. (Arrows go from *implicans* to *implicatum*.)

These implications suffice to decide the question of whether the five logicist theses are true, in the context of the first-order view. For, assuming paraphrase-functions are minimally adequate, it is a consequence of Gödel's Incompleteness Theorem that the syntactic version of *Consequence-Logicism* is false for any interesting fragment of mathematics.[9] And, since each of the other logicist theses implies the syntactic version of *Consequence-Logicism*, it follows that each of the other logicist theses must also be false.

2.3 The Higher-order View

When the logicist theses are read in accordance with the higher-order view, the following relations obtain: (see figure 2)

[HOV-1] The syntactic versions of *Truth-Logicism* and *Consequence-Logicism* imply the corresponding semantic versions, since higher-order logic is sound according to the higher-order view. (The converse implications do not hold because higher-order logic is incomplete.)

[HOV-2] Each version of *Truth-Logicism* implies the corresponding version of

---

then every sentence in T follows from a sentence stating that there are precisely n objects, for some n; if T has an infinite model, then every sentence in T follows from a set of sentences to the effect that the universe is infinite.

[9] *Proof Sketch:* Let T be a sufficiently interesting mathematical theory (i.e. a theory in which Robinson Arithmetic can be interpreted), and suppose there is a minimally-adequate paraphrase-function $*$ which delivers the syntactic version of *Consequence-Logicism* with respect to a recursive axiomatization $\mathcal{A}$ of T. Then there is a decision procedure for Robinson Arithmetic. [For ϕ a sentence of Robinson Arithmetic, let ψ be the interpretation of ϕ in T. Since $*$ is minimally adequate, it is recursive. And, since there is a recursive function listing the theorems derivable from $\mathcal{A}$, the set S of syntactic consequences of $\mathcal{A}$ is recursive. But, since *Consequence-Logicism* holds, ψ^* is in S if ψ is true, and the negation of ψ^* is in S if ψ is false.] This contradicts Gödel's Incompleteness Theorem, which implies that that there is no decision procedure for Robinson Arithmetic.

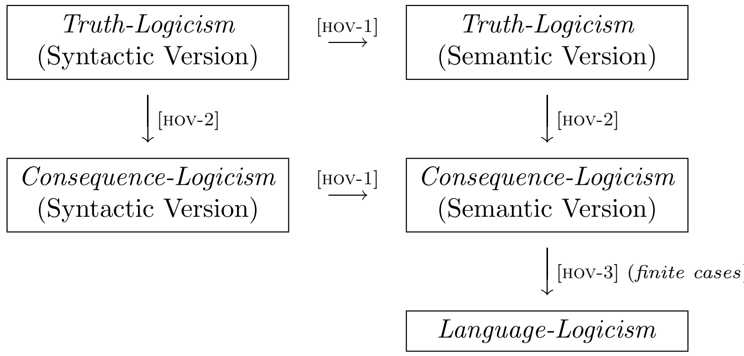

Figure 2: Implications amongst the five logicist theses, when read in accordance with the higher-order view. (Arrows go from *implicans* to *implicatum.*)

Consequence-Logicism, since *Truth-Logicism* is the special case of *Consequence-Logicism* when $\mathcal{A}$ is empty.

[HOV-3] If the semantic version of *Consequence-Logicism* obtains on the basis of a finite axiomatization, then there is a minimally adequate paraphrase-function which makes *Language-Logicism* true.[10]

The final implication holds because, when the semantic version of *Consequence-Logicism* obtains on the basis of a finite axiomatization, $\mathcal{A}$, the paraphrase-function taking a higher-order sentence $\ulcorner \phi \urcorner$ to the result of universally Ramsifying $\ulcorner \mathcal{A} \to \phi \urcorner$ is minimally adequate.[11] And the result of universally Ramsifying $\ulcorner \mathcal{A}^c \to \phi \urcorner$ is a higher-order sentence containing no non-logical vocabulary.

Let us now turn to an assessment of the five logicist theses, from the perspective of the higher-order view.

2.3.1 The Syntactic Versions of *Consequence-* and *Truth-Logicism*

When it comes to the syntactic versions of *Consequence-* and *Truth-Logicism*, the higher-order view constitutes no improvement over the first-order view. Assuming paraphrase-functions are minimally adequate, it is a consequence of Gödel's Incompleteness Theorem that the syntactic version of *Consequence-Logicism* is false for any interesting fragment of mathematics.[12] And, since the syntactic

[10] It is worth emphasizing that, whereas [FOV-3] describes an implication from *Language-Logicism* to *Consequence-Logicism*, [HOV-3] describes an implication from *Consequence-Logicism* to *Language-Logicism*. Our proof of [FOV-3] cannot be reproduced in the context of the higher-order view because it makes use of the Downward Löwenheim-Skolem Theorem, which holds for first-order languages but not their higher-order counterparts. Our proof of [HOV-3] cannot be reproduced in the context of the first-order view because, although the result of Ramsifying a higher-order sentence is always a higher-order sentence, the result of Ramsifying a first-order sentence is not always a first-order sentence.

[11] The result of *universally Ramsifying* $\ulcorner \mathcal{A}^c \to \phi \urcorner$ is the universal closure of the result of uniformly substituting variables of the appropriate type for all non-logical primitives in $\ulcorner \mathcal{A}^c \to \phi \urcorner$.

[12] The proof is analogous to that in footnote 9.

version of *Truth-Logicism* implies the syntactic version of *Consequence-Logicism*, it follows that the syntactic version of *Truth-Logicism* must also be false.

2.3.2 The Semantic Version of *Consequence-Logicism*

Because there are semantic consequences of higher-order theories which are not derivable on the basis of any sound deductive system, nothing we have said so far entails that the *semantic* version of *Consequence-Logicism* must fail in the context of the higher-order view. But does it?

For the special case of arithmetic, we have a positive result, thanks to semantic completeness: the second-order Dedekind-Peano Axioms semantically imply every true sentence of the language of pure second-order arithmetic.[13] From this the semantic version of *Consequence-Logicism* follows immediately, in the context of the higher-order view.

Similar results hold for other branches of pure mathematics.[14] In the case of set theory, however, there is a problem. Unlike the second-order Dedekind-Peano Axioms, the axioms of second-order ZFC are not guaranteed to be semantically complete: there may be a sentence of the language of pure second-order set theory such that neither it nor its negation is a semantic consequence of the axioms.[15] All we have is a partial result: if we add to the axioms of second-order ZFC a sentence specifying how many sets there are, then we are guaranteed semantic completeness.[16]

Thus, one possibility is to enrich the axioms of second-order ZFC with a sentence S_μ of pure second-order logic to the effect that there are no inaccessible cardinals. That would certainly yield semantic completeness. Unfortunately, S_μ is not universally regarded as true. Many set-theorists endorse a Global Reflection Principle, to the effect that there is a set α such that every true sentence in the language of higher-order set-theory is true when the quantifiers are restricted to the elements of α. And the Global Reflection Principle implies that S_μ is false because it implies that there is an inaccessible cardinal (since the standard set theoretic axioms are true, the Reflection Principle implies that they are also true when the quantifiers are restricted to the members of some set, and this is only possible if there is an inaccessible cardinal.)

As it turns out, matters are worse still: if the Global Reflection Principle is true, then no recursive set of true sentences in the language of higher-order set theory semantically implies every true sentence in the language of higher-order set-theory.[17] So the Global Reflection Principle implies that *there is no hope* of producing a semantic completeness result in the case of set theory.

[13]The original result is due to Dedekind. For a proof see Shapiro (1991), Theorem 4.8.

[14]For a proof of the categoricity of analysis, see Shapiro (1991), Theorem 4.10.

[15]This will be the case if there are models of second-order set theory whose domain is larger than the first strongly inaccessible cardinal.

[16]The result is due to Zermelo. See Zermelo (1930).

[17]*Proof Sketch:* Suppose, for *reductio*, that $\mathcal{A}$ is such a set. Define, within the language of higher-order set-theory, the notion of a (set-sized) model for the language of higher-order set-theory, and a corresponding satisfaction predicate. Let $\ulcorner \bar{\mathcal{S}} \models^+ \bar{\phi} \urcorner$ be a sentence of the language of higher-order set theory stating that every (set-sized) model in which every sentence whose Gödel number is in $\bar{\mathcal{S}}$ is satisfiable is also a model in which the sentence with Gödel-

So much for pure mathematics. In the case of *applied* mathematics it is unreasonable to ask for a semantic completeness result, since no set of mathematical axioms should imply, for instance, that Mars has two moons (or that it doesn't). But one might hope for a *relative* semantic completeness result, to the effect that there is a set of axioms which, together with the set of true sentences containing no mathematical vocabulary, semantically implies every true sentence in some particular fragment of applied mathematics, and the negation of every false one. As it turns out, it is possible to prove a relative completeness result for higher-order applied arithmetic.[18] So *Consequence-Logicism* holds for higher-order applied arithmetic, *modulo* the set of non-arithmetical truths.

2.3.3 *Language-Logicism*

When all we require of paraphrase-functions is minimal adequacy, it follows from [HOV-3] that any higher-order theory with a finite and complete axiomatization can be paraphrased using no non-logical vocabulary. So, thanks to the semantic completeness results discussed in the preceding section, *Language-Logicism* can be shown to be true for the case of pure higher-order arithmetic.

Of course, minimal adequacy is a very weak constraint on paraphrase-functions, and it is likely to leave some advocates of logicism unsatisfied. One way to strengthen the constraint is by making use of semantic notions. For instance, one might require paraphrase-functions to preserve *content*. We will have something to say about this in section 3, but for the moment it is best to focus on a syntactic constraint instead. Informally, say that a paraphrase-function $*$ is *compositional* just in case the compositional structure of ϕ^* mirrors the compositional structure of ϕ, for every ϕ in the domain of $*$. (For a formal characterization of this constraint, see the appendix.) Paraphrase-functions based on Ramsey-conditionals are not compositional. So, when compositionality is required, Ramsey-conditionals do not deliver *Language-Logicism*.

For the case of pure higher-order arithmetic, however, there are compositional paraphrase-functions which deliver formulas of pure higher-order logic as outputs. One is implicit in Whitehead and Russell's *Principia* (see 'Logicism', in this volume). More recently, Harold Hodes has set forth a paraphrase-function that works by treating first-order quantifiers ranging over natural numbers as *third-order* quantifiers ranging over finite cardinality object-quantifiers (that is, second-order concepts true of all and only first-order concepts under which precisely n objects fall, for some n).[19] Hodes's paraphrase-function can be shown

number $\bar{\phi}$ is satisfiable. Since $\mathcal{A}$ is recursive, '$x \in \mathcal{A}$' is expressible in second-order ZFC. So $\ulcorner \bar{\mathcal{A}} \models^+ \bar{\phi} \urcorner$ is a truth predicate for the language of higher-order set-theory. [By our assumption, every true sentence of the language of higher-order set-theory is a semantic consequence of $\mathcal{A}$, and hence true on every (set-sized) model; and, by the Global Reflection Principle, there is a (set-sized) model according to which every untrue sentence of the language of higher-order arithmetic is false.] But it follows from Tarski's Theorem that no language can express its own truth-predicate.

[18]See Rayo (forthcoming). It is worth noting that, whereas completeness results are usually based on categoricity theorems, the proof in Rayo (forthcoming) relies on an essentially different method.

[19]See Hodes (1984) and Hodes (1990). See also Wright (1983) pp. 36-40, Bostock (1979)

to be minimally adequate on the assumption that the universe is infinite, and its restriction to pure higher-order arithmetic always delivers formulas of pure higher-order logic as outputs. So, when no more is required of paraphrase-functions than compositionality and minimal adequacy, *Language-Logicism* can be shown to be true for the case of pure higher-order arithmetic in the context of the higher-order view.

2.3.4 The Semantic Version of *Truth-Logicism*

We noted in section 2.3.1 that the syntactic version of *Truth-Logicism* fails in the context of the higher-order view. But, for all that has been said so far, it might still be the case that the *semantic* version of *Truth-Logicism* holds.

None of the paraphrase-functions we have considered so far deliver this result. Consider, for instance, the paraphrase-function $\Re$, which takes each sentence $\ulcorner \phi \urcorner$ of pure second-order arithmetic to the result of universally Ramsifying $\ulcorner \mathcal{A} \to \phi \urcorner$ (where $\ulcorner \mathcal{A} \urcorner$ is the conjunction of the Dedekind-Peano Axioms). In order for $\Re$ to deliver *Truth-Logicism*, $\phi^{\Re}$ would have to be a logical truth whenever ϕ is true on the intended interpretation, and a logical falsehood whenever ϕ is false on the intended interpretation. This is certainly the case when we restrict our attention to models with infinite domains. But it is not the case when models with finite domains are allowed, since $\phi^{\Re}$ is true on any model with a finite domain, independently of whether ϕ is true on the intended interpretation.

A similar problem afflicts Hodes's paraphrase-function. When we restrict our attention to models with infinite domains, the Hodes-paraphrase of ϕ is a logical truth just in case ϕ is true on the intended interpretation, for any sentence ϕ in the language of higher-order arithmetic. But this is not the case when models with finite domains are allowed.

Might other paraphrase-functions do better? In an important sense, the answer is 'no'. On the assumption that paraphrase-functions are required to be compositional, the semantic version of *Truth-Logicism* must fail for any interesting fragment of mathematics, in the context of the higher-order view. (See appendix for proof.)

A final point is worth mentioning. So far we have made the (standard) assumption that first- and higher-order quantifiers are *domain-relative*. In other words, we have presupposed that which individuals the truth-value of a quantified sentence depends on is not a logical matter and, accordingly, that the range of the quantifiers may vary from model to model. But Timothy Williamson has recently made a case for *logically unrestricted* quantifiers: quantifiers whose range consists, as a matter of logic, of absolutely everything.[20] If Williamson's logically unrestricted quantifiers are allowed, then the fact that there are at least two objects in existence suffices to guarantee that, when the quantifiers are regarded as logically unrestricted, '$\exists x \exists y (x \neq y)$' is logically true. Similarly, the fact that there are infinitely many objects in existence suffices to guarantee that, when the quantifiers are regarded as logically unrestricted, a second-

volume II chapter 1, and Rayo (forthcoming).

[20] See Williamson (1999) and Rayo and Williamson (forthcoming).

order sentence stating that the universe is infinite is logically true. And, of course, when logically unrestricted quantifiers are allowed, the status of *Truth-Logicism* changes dramatically. The fact that there are infinitely many objects ensures that, when quantifiers are regarded as logically unrestricted, the Hodes-paraphrase of ϕ is a logical truth (falsehood) just in case ϕ is true (false) on the intended interpretation, for any ϕ in the language of pure higher-order arithmetic. It also ensures that, when quantifiers are regarded as logically unrestricted, $\phi^{\Re}$ is a logical truth (falsehood) whenever ϕ is true (false) on the intended interpretation, for any ϕ in the language of pure second-order arithmetic. Finally, when the quantifiers are regarded as logically unrestricted, one can show that there is a minimally adequate (but non-compositional) paraphrase-function verifying the semantic version of *Truth-Logicism* for the language of pure higher-order set theory, given the assumption that there are strongly inaccessibly many objects.[21]

This concludes our assessment of the five logicist theses. Figure 3 summarizes some of the main results.

2.4 What Have We Learned?

We have seen that *Truth-Logicism* fails on natural assumptions about logic and paraphrase. Although this casts considerable doubt on the view that every mathematical truth is a truth of logic, it does not constitute a decisive refutation. One way of maintaining *Truth-Logicism* is by adopting a conception of logic distinct from the first- and higher-order views. Frege, for example, took Basic Law V—the principle that the extension of the Fs equals the extension of the Gs just in case the Fs are the Gs—to be 'purely logical'.[22] (Though, of course, he later discovered that Basic Law V leads to contradiction.) Another way of resisting the first- and higher-order views is by adopting Williamson's logically unrestricted quantifiers. But it is important to be clear that logically unrestricted quantifiers have an epistemological cost, since their legitimacy would destroy any hope of establishing a straightforward connection between logical truth and *a priori* knowledge.[23] The failure of *Truth-Logicism* on the first- and higher-order views makes it seem implausible that one could acquire *a priori* knowledge of the standard mathematical axioms merely by carrying out logical derivations. (We will see in section 4 that Neo-Fregeans have argued for the weaker claim that one can acquire *a priori* knowledge of the standard arith-

[21] Consider the paraphrase-function $\circledast$, which takes each sentence $\ulcorner\phi\urcorner$ of the language of higher-order set theory with urelements to the result of universally Ramsifying $\ulcorner$ZFCU $\rightarrow \phi\urcorner$ (where $\ulcorner$ZFCU$\urcorner$ is the conjunction of the axioms of second-order ZFC with urelements plus an axiom to the effect that the urelements for a set). If there are strongly inaccessibly many objects, it follows from the result in McGee (1997) that, when quantifiers are regarded as logically unrestricted, $\phi^{\circledast}$ is a logical truth (falsehood) whenever ϕ is true (false) on the intended interpretation for every ϕ in the langauge of *pure* higher-order arithmetic.

[22] See Frege (1893 and 1903), vol 1, p. vii.

[23] See, however, Williamson (1999). It is also worth noting that, because of Gödel's incompleteness results, the legitimacy of higher-order quantification should be enough to make us suspicious of a straightforward connection between logical truth and *a priori* knowledge.

	First-order View allowing only minimally adequate paraphrase-functions.	Second-order View allowing all and only compositional, minimally adequate paraphrase-functions.
Language-Logicism	False	True$^{\alpha}$
Consequence-Logicism (Semantic Version)	False	True$^{\beta}$
Consequence-Logicism (Syntactic Version)	False	False
Truth-Logicism (Semantic Version)	False	False$^{\gamma}$
Truth-Logicism (Syntactic Version)	False	False

$^{\alpha}$ for the case of pure higher-order arithmetic.

$^{\beta}$ for the case of pure and applied higher-order arithmetic (*modulo* the set of non-arithmetical truths).

$^{\gamma}$ true for the case of pure higher-order arithmetic when logically unrestricted quantifiers are allowed.

Figure 3: Summary

metical axioms merely by setting forth an appropriate stipulation and going on to carry out appropriate logical derivations.)

We have also seen that, on the higher-order view, the semantic version of *Consequence-Logicism* holds for the case of (pure and applied) higher-order arithmetic. This delivers the result that every arithmetical sentence is either determinately true or determinately false, provided only that appropriate arithmetical axioms are determinately true and that higher-order consequence preserves determinate truth.[24] As noted in section 2.3.2, the Global Reflection Principle implies that there is no hope of producing a semantic completeness result for the case of set theory, but it is worth noting that it does not preclude a determinacy result. By working within the language of set theory with urelements, McGee (1997) and Uzquiano (2002) have found axiomatizations of second-order set theory with the feature that the pure sets of any two models are isomorphic, provided only that the models in question have domains consisting of everything there is.[25] Accordingly, by insisting that one's quantifiers are to range over absolutely everything,[26] one could carry out a recursively specifiable

[24] All of this, *modulo* indeterminacy in non-mathematical expressions.

[25] For discussion on quantifying over everything see Parsons (1974), Dummett (1981) chapters 14-16, Cartwright (1994), Boolos (1998b), Williamson (1999), McGee (2000), the postscript to Field (1998) in Field (2001), Rayo (2002), Rayo (2003b), Rayo and Williamson (forthcoming), Glanzberg (typescript), and Williamson (typescript).

[26] The possibility of doing so determinately is defended in McGee (2000) and Rayo (2003b).

stipulation capable of delivering the result that every sentence in the language of pure set-theory is either determinately true or determinately false. As before, this relies on the assumption that the appropriate set-theoretic axioms are determinately true and that higher-order semantic consequence preserves determinate truth.

Finally, we have seen that, on the higher-order view, *Language-Logicism* holds for the case of pure higher-order arithmetic. This result might be used as part of a nominalist account of mathematics. (And, of course, nominalism is of interest even if one is not a nominalist; for instance, a nominalist account of higher-order applied arithmetic would conclusively undermine the thought that we are justified in believing standard mathematical theories because they are indispensable for the natural sciences.)

3 Recarving Contents

The only *semantic* constraint on paraphrase-functions we have considered so far is preservation of truth-value. Are there other semantic constraints one might reasonably impose?

One option, of course, is to require strict synonymy. But perhaps a more liberal constraint can be extracted from the following passage in Frege's *Foundations of Arithmetic*:

> The judgement 'line a is parallel to line b', or, using symbols, '$a // b$', can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'the direction of line a is identical with the direction of line b'. Thus we replace the symbol // by the more generic symbol =, through removing what is specific in the content of the former and dividing it between a and b. We carve up the content in a way different from the original way, and this yields a new concept (§64).

Thus, the sentences 'line a is parallel to line b' and 'the direction of line a is identical to the direction of line b' are taken to express the same content, only 'carved up' in a different way. And, as Frege observed, something similar might be said for the case of *arithmetic*: the sentences 'the Fs can be put in one-one correspondence with the Gs' and 'the number of the Fs is identical to the number of the Gs' might be taken express the same content, only 'carved up' in a different way.

In this section we will consider the following three questions:

1. Is there an interesting way of making the notion of content-recarving precise?

2. Is there a paraphrase-function $\mathcal{I}$ such that, for any sentence ϕ in some interesting fragment of mathematics, ϕ and $\phi^{\mathcal{I}}$ express the same content, only 'carved up' in a different way?

3. If $\mathcal{I}$ exists, does it verify any of the logicist theses?

It is best for expositional purposes to begin with question 2, while working with the informal characterization of content-recarving provided by Frege's remark. Later we will return to question 1, and say something about how the notion of content-recarving might be made precise. We end with question 3.

3.1 The Second Question

First, some notation. We let '$F \approx G$' be a second-order formula expressing one-one correspondence between the Fs and the Gs,[27] and interpret '$\mathbf{N}(n, F)$' as 'n numbers the Fs'.[28]

Next, we set forth a paraphrase-function $\mathcal{F}$. Intuitively, $\mathcal{F}$ works by paraphrasing talk of the *number* of the Fs by talk of the Fs themselves.[29] Formally, $\mathcal{F}$ may be characterized as follows,where 'm_1', 'm_2', ... are first-order variables ranging over numbers, and 'Z_1', 'Z_2', ... are second-order variables ranging over an unrestricted domain:

- $\ulcorner \forall m_i(\phi) \urcorner^{\mathcal{F}} \; = \; \ulcorner \forall Z_i \urcorner \frown \ulcorner \phi \urcorner^{\mathcal{F}}$;
- $\ulcorner m_i = m_j \urcorner^{\mathcal{F}} \; = \; \ulcorner Z_i \approx Z_j \urcorner$;
- $\ulcorner \mathbf{N}(m_i, X) \urcorner^{\mathcal{F}} \; = \; \ulcorner Z_i \approx X \urcorner$;
- $\ulcorner \neg\phi \urcorner^{\mathcal{F}} \; = \; `\neg' \frown \ulcorner \phi \urcorner^{\mathcal{F}}$;
- $\ulcorner \phi \wedge \psi \urcorner^{\mathcal{F}} \; = \; \ulcorner \phi \urcorner^{\mathcal{F}} \frown `\wedge' \frown \ulcorner \psi \urcorner^{\mathcal{F}}$;
- $\ulcorner \phi \rightarrow \psi \urcorner^{\mathcal{F}} \; = \; \ulcorner \phi \urcorner^{\mathcal{F}} \frown `\rightarrow' \frown \ulcorner \psi \urcorner^{\mathcal{F}}$.

As an example, consider the sentence 'the number of the Fs = the number of the Gs'. In our notation, it can be formulated as:

$$\forall m_1 \forall m_2 (\mathbf{N}(m_1, F) \wedge \mathbf{N}(m_2, G) \rightarrow m_1 = m_2).$$

The result of applying $\mathcal{F}$ is:

$$\forall Z_1 \forall Z_2 (Z_1 \approx F \wedge Z_2 \approx G \rightarrow Z_1 \approx Z_2),$$

which is equivalent to:

$$F \approx G.$$

[27] That is,

$$F \approx G \;\; \equiv_{df} \;\; \exists R \; [\forall w(Fw \rightarrow \exists! v(Gv \wedge Rwv)) \wedge \forall w(Gw \rightarrow \exists! v(Fv \wedge Rvw))].$$

[28] '$\mathbf{N}$' is a second-level predicate because it takes a second-order variable in one of its argument-places. For a discussion of second-level predicates see Rayo (2002).

[29] I use plurals for expository purposes only. I do not presuppose that second-order quantifiers should be understood plurally, as in Boolos (1984).

Thus, the result of applying $\mathcal{F}$ to 'the number of the Fs = the number of the Gs' is equivalent to '$F \approx G$'. In a Fregean spirit, one might be tempted to say that the application of $\mathcal{F}$ has resulted in a recarving of content.

What happens when we apply $\mathcal{F}$ to other arithmetical formulas? For instance, what happens when we apply it to a sentence such as 'the number of the planets = 9'? When numerical predicates are defined in the obvious way,[30]

$$\begin{aligned} 0(m) &\equiv_{df} \forall X(\mathbf{N}(m,X) \leftrightarrow \neg\exists x(X(x))); \\ 1(m) &\equiv_{df} \forall X(\mathbf{N}(m,X) \leftrightarrow \exists!_1 x(X(x))); \\ 2(m) &\equiv_{df} \forall X(\mathbf{N}(m,X) \leftrightarrow \exists!_2 x(X(x))); \\ &\ldots \end{aligned}$$

'the number of the planets = 9' can be formulated as:

$$\forall m_1(\mathbf{N}(m_1, \hat{x}[\textsc{Planet}(x)]) \rightarrow 9(m_1)).^{31}$$

And the result of applying $\mathcal{F}$ to this sentence is equivalent to:

$$\forall Z_1(Z_1 \approx \hat{x}[\textsc{Planet}(x)] \rightarrow \exists!_9 x(Z_1(x))),$$

which is in turn equivalent to:

$$\exists!_9 x(\textsc{Planet}(x)).$$

Thus, the result of applying $\mathcal{F}$ to 'the number of the planets = 9' is equivalent to '$\exists!_9 x(\text{Planet}(x))$'. One might again be tempted to say, in a Fregean spirit, that the application of $\mathcal{F}$ has resulted in a recarving of content.

It is easy to define 'Plus' and 'Times' in terms of $\mathbf{N}(n, F)$. Consider 'Plus' as an example:

$$\begin{aligned} &\textsc{Plus}(m_1, m_2, m_3) \equiv_{df} \\ &\qquad \forall X \forall Y \forall Z(\mathbf{N}(m_1, X) \wedge \mathbf{N}(m_2, Y) \wedge \mathbf{N}(m_3, Z) \rightarrow \textsc{Join}(X, Y, Z)); \end{aligned}$$

where

$$\begin{aligned} &\textsc{Join}(X, Y, Z) \equiv_{df} \\ &\qquad \exists X'(X' \approx X \wedge \forall x(X'(x) \rightarrow \neg Y(x)) \wedge \hat{x}[X'(x) \vee Y(x)] \approx Z). \end{aligned}$$

[30] As usual,

$$\begin{aligned} \exists!_1 x(\phi(x)) &\equiv_{df} \exists x(\phi(x) \wedge \neg\exists y(y \neq x \wedge \phi(y))), \\ \exists!_{n+1} x(\phi(x)) &\equiv_{df} \exists x(\phi(x) \wedge \exists!_n y(y \neq x \wedge \phi(y))). \end{aligned}$$

[31] Syntactically, an expression of the form '$\hat{x}\ [\phi(x)]$' takes the place of a monadic second-order variable. But the result of substituting '$\hat{x}\ [\phi(x)]$' for 'Y' in a formula '$\Psi(Y)$' is to be understood as shorthand for:

$$\forall W\ (\forall x(W\,x \leftrightarrow \phi(x)) \rightarrow \Psi(W)).$$

Additional clauses must be added to the definition of $\mathcal{F}$ before it can be applied to sentences containing '$\hat{x}\ [\phi(x)]$', but they are all trivial. See Rayo (forthcoming) for details.

A sentence such as '$\forall n \forall m(n+m=m+n)$' can then be formulated as:

$$\forall m_1 \forall m_2 \forall m_3(\textsc{Plus}(m_1, m_2, m_3) \rightarrow \textsc{Plus}(m_2, m_1, m_3)).$$

The result of applying $\mathcal{F}$ is equivalent to:[32]

$$\forall Z_1 \forall Z_2 \forall Z_3(\textsc{Join}(Z_1, Z_2, Z_3) \rightarrow \textsc{Join}(Z_2, Z_1, Z_3)).$$

And, again, in a Fregean spirit, one might be tempted to say that the application of $\mathcal{F}$ has resulted in a recarving of content.

As it turns out, $\mathcal{F}$ can be made to encompass every sentence of pure and applied nth-order arithmetic, and it can be done in such a way that the result of applying $\mathcal{F}$ to a sentence of pure n-th order arithmetic is always a sentence of pure second-order logic. Moreover, as one might have hoped, $\mathcal{F}$ is a compositional and minimally adequate paraphrase-function.[33]

3.2 The First Question

Is there any interesting way of making the notion of content-recarving precise? Hale (1997) sets forth one proposal, which is intended as a defence of the Neo-Fregean Program. Here we will consider another, which is not. It is based on the following definition:

> Let A and B be sets of formulas, let $\mathcal{I}$ be a one-one function from A onto B, and suppose that the following conditions are met:
>
> 1. $\mathcal{I}$ preserves the *logical networking* of A.
> 2. For any $\phi \in A$, ϕ and $\phi^{\mathcal{I}}$ are *equisatisfiable*.
>
> Then we shall say that, for any $\phi \in A$, ϕ is an $AB_{\mathcal{I}}$-recarving of $\phi^{\mathcal{I}}$.

Although the appendix provides a precise characterization of the notions of preservation of logical networking and equisatisfiability, an example should suffice to convey the intuitive idea.

Let A be the set of arithmetical formulas,[34] and let B be the result of applying $\mathcal{F}$ to formulas in A. $\mathcal{F}$ preserves the logical networking of A because

[32] Additional clauses must be added to the definition of $\mathcal{F}$ before it can be applied to this sentence, but they are all trivial. See Rayo (forthcoming) for details.

[33] See Rayo (forthcoming).

[34] More specifically, A is the set of formulas of a second-order language L, containing the following kinds of variables: first-order arithmetical variables, 'm_1', 'm_2', ..., first-order general variables, 'x_1', 'x_2', ..., and, for n a positive integer, n-place second-order general variables $\ulcorner X_1^n \urcorner$, $\ulcorner X_2^n \urcorner$, ... To avoid variable clashes, monadic second-order general variables are divided in two groups: the $\ulcorner X_{2i}^1 \urcorner$—which are abbreviated $\ulcorner Z_i \urcorner$—are associated with first-order arithmetical variables by $\mathcal{F}$; the $\ulcorner X_{2i+1}^1 \urcorner$—which are abbreviated $\ulcorner X_i \urcorner$—are used for more general purposes. We assume that L has been enriched with a single higher-level predicate '$\mathbf{N}$' taking a first-order arithmetical variable in its first argument-place, and a monadic second-order general variable of the second group in its second argument-place. The well-formed formulas of L are defined in the usual way, with the proviso that an atomic formula can contain arithmetical variables only if it is of the form $\ulcorner m_i = m_j \urcorner$ or $\ulcorner \mathbf{N}(m_i, X_j) \urcorner$.

it satisfies three conditions. First, $\mathcal{F}$ is compositional. Second, $\mathcal{F}$ respects the logical connectives of sentences in A (e.g., $\ulcorner \neg\phi \urcorner^{\mathcal{F}} =$ '$\neg$' $\frown \ulcorner \phi \urcorner^{\mathcal{F}}$). Third, $\mathcal{F}$ maps identity-statements in A to formulas that 'function' as identity statements in B. For instance, $\mathcal{F}$ maps '$m_1 = m_2$' to '$Z_1 \approx Z_2$', which 'functions' as an identity statement in B because (i) '$\approx$' is reflexive, symmetric and transitive, and (ii) for any sentence $\ulcorner \phi \urcorner$ of B containing only $\ulcorner Z_i \urcorner$ free, the universal closure of

$$Z_i \approx Z_j \rightarrow (\phi \leftrightarrow \phi[Z_j/Z_i])$$

is true.[35] (It is important to note that, although '$Z_1 \approx Z_2$' 'functions' as an identity statement within the context of B, it doesn't do so within the context of the entire language, since condition (ii) does not generally hold for sentences outside B. For instance, the universal closure of '$Z_1 \approx Z_2 \rightarrow (Z_1(a) \leftrightarrow Z_2(a))$' is always false.)

Let us now turn to equisatisfiability. Consider the following sentence, which is sometimes called 'HP':

$$\forall F \forall G(\hat{m}[\mathbf{N}(m, F)] = \hat{m}[\mathbf{N}(m, G)] \leftrightarrow F \approx G).$$[36]

Assuming HP, ϕ and $\phi^{\mathcal{F}}$ can be shown to be equisatisfiable for any $\phi \in A$. In essence, this is because two conditions are met. First, $\mathcal{F}$ is compositional. Second, we have the following:

> Let $\phi(m_i)$ be a formula in A containing only $\ulcorner m_i \urcorner$ free, and let $\psi(Z_i)$ be the result of applying $\mathcal{F}$ to $\phi(m_i)$. Then $\phi(m_i)$ is true of the *number* of the Gs just in case $\psi(Z_i)$ is true of the Gs.

Putting all of the above together, we may conclude that, on the assumption that HP is true, ϕ is an $AB_{\mathcal{F}}$-recarving of $\phi^{\mathcal{F}}$ for any ϕ in A. A similar result can be proved for Hodes's paraphrase-function, described in section 2.3.3.

Now, in what sense do $AB_{\mathcal{I}}$-recarvings provide us with an elucidation of the intuitive notion a content-recarving? On the present view, it makes no sense to ask whether ϕ expresses a recarving of the content expressed by ϕ' when ϕ and ϕ' are considered in isolation. Rather, one must ask whether ϕ expresses a recarving of the content expressed by ϕ when ϕ is considered in the context of a set of sentences A and ϕ' is considered in the context of a set of sentences B. The proposal is then this:

[35] '$\phi[Z_j/Z_i]$' is the result of substituting $\ulcorner Z_j \urcorner$ for $\ulcorner Z_i \urcorner$ in $\ulcorner \phi \urcorner$, and possibly relabelling bound variables to avoid clashes.

[36] When HP is formulated within a one-sorted higher-order language it entails that the universe is infinite. But when it is formulated in a two-sorted higher-order language (such as the language described in footnote 35), it is comptaible with a finite universe (it entails only that *either* the range of the arithmetical variables and the range of the general variables are both infinite, *or* there is one more object in the range of the arithmetical variables than in the range of the general variables). It is worth noting that the equisatisfiability result for $\mathcal{F}$ requires no more than the two-sorted reading of HP.

In the context of our discussion of the Neo-Fregean Program in section 4, however, it is crucial that HP be read as formulated within a one-sorted higher-order language. This is because, unlike its two-sorted counterpart, the one-sorted version of HP entails a version of the Dedekind-Peano Axioms. For more on formulations of HP within multi-sorted languages, see Heck (1997b).

> The content expressed by ϕ is a recarving of the content expressed by ψ (when ϕ is considered in the context of a set of sentences A and ϕ' is considered in the context of a set of sentences B) just in case the context expressed by ϕ *relative to* A is the same the content expressed by ψ *relative to* B.

What is it to express a content relative to a set of sentences? Here is an informal explanation. Suppose it is right to say that ϕ and ψ express the same content *simpliciter* just in case: (i) ϕ and ψ have the same logical structure; and (ii) non-logical primitives occupying corresponding places in the logical structure of ϕ and ψ have the same semantic-value. One can then say that the content expressed by ϕ relative to A is the same as the content expressed by ψ relative to B just in case: (i) the logical structure, LS^{ϕ}, that is 'displayed' by ϕ in the context of A is isomorphic to the logical structure, LS^{ψ}, that that is 'displayed' by ψ in the context of B; and (ii) if E is an expression in ϕ that 'functions' in the context of A as a non-logical primitive and occupies a certain position in LS^{ϕ}, and E' is an expression that 'functions' in the context of B as a non-logical primitive and occupies a corresponding position in LS^{ψ}, then the semantic-value of E is 'analogous' to the semantic-value of E'.

The notion of an $AB_{\mathcal{I}}$-recarving embodies a rigorous version of this informal explanation. The requirement of preservation of logical networking is a way of capturing the thought that the logical structure 'displayed' by a sentence ϕ in A is isomorphic to the logical structure 'displayed' by $\phi^{\mathcal{I}}$ in B. And the requirement of equisatisfiability is a way of capturing the thought that any expression in ϕ that 'functions' in the context of A as a non-logical primitive is mapped onto an expression of $\phi^{\mathcal{I}}$ that 'functions' in the context of B as non-logical primitive with an 'analogous' semantic-value.

3.3 The Third Question

What can be concluded about the logicist theses? Since the result of applying $\mathcal{F}$ to a sentence of pure n-th order arithmetic is always a sentence of pure second-order logic, $\mathcal{F}$ can be shown to verify *Language-Logicism*, in the context of the higher-order view. In addition, $\mathcal{F}$ can be shown to verify the semantic version of *Consequence-Logicism* for the case of pure and applied nth-order arithmetic.[37,38] When the intuitive notion of content-recarving is cashed out in terms of $AB_{\mathcal{F}}$-recarvings and HP is assumed, this yields the result that *Language-* and *Consequence-Logicism* are both verified by a paraphrase-function which preserves *content*, modulo recarvings.

[37] This follows from the completeness result in Rayo (forthcoming).

[38] Since $\mathcal{F}$ is a compositional paraphrase-function, the results in section 2.3 imply that it cannot verify any of the three remaining logicist theses. However, if logically unrestricted quantifiers are allowed, then $\mathcal{F}$ verifies the semantic version of *Truth-Logicism*.

4 The Neo-Fregean Program

This final section will be devoted to the Neo-Fregean Program (or 'Neo-Logicism', as it is sometimes called). Despite its roots as an attempt to rescue Frege's logicist project from inconsistency, Neo-Fregeanism has increasingly developed a life of its own, and must be assessed on its own terms.[39] This makes the present section largely independent from the rest of the paper.

The core of the Neo-Fregean Program is the contention that it is possible to acquire an *a priori* justification for arithmetical statements in a special kind of way. Let $\langle \mathbf{N}, \mathrm{HP} \rangle$ be the following linguistic stipulation:

> The second-level predicate '**N**' is to be used in such a way that HP (introduced in section 3.2) turns out to be true.

Neo-Fregeans believe that, merely as a result of setting forth $\langle \mathbf{N}, \mathrm{HP} \rangle$, one can acquire an *a priori* justification for the belief that HP is true. But HP deductively implies (definitional equivalents of) the second-order Dedekind-Peano Axioms.[40] So, if the Neo-Fregean story is right, then $\langle \mathbf{N}, \mathrm{HP} \rangle$ yields an *a priori* justification for the second-order Dedekind-Peano axioms.[41]

The Neo-Fregean Program has given rise to many objections and replies. We will make no attempt to survey the literature here.[42] Instead, we will focus on a problem which I believe is especially important, and surprisingly underdeveloped. (For a broader perspective on the Neo-Fregean Program, see 'Logicism in the 21st Century', in this volume.)

The problem arises as follows. Say that a linguistic stipulation is *weakly successful* if it has the effect of rendering its definienda meaningful. Say that a linguistic stipulation is *strongly successful* if it has the effect of rendering its definienda meaningful in such a way that the sentence targeted by the stipulation turns out to be true. $\langle \mathbf{N}, \mathrm{HP} \rangle$ is of little interest to the Neo-Fregean unless it is strongly successful. (If it is only weakly successful, then it doesn't have the effect of rendering HP true, and *a fortiori* the Neo-Fregean's proposal cannot yield the result that HP is *known* to be true.) But HP can only be true if there are infinitely many objects.[43] So $\langle \mathbf{N}, \mathrm{HP} \rangle$ can only be strongly successful if there are infinitely many objects.

So far, all we have is the conclusion that either there are infinitely many objects or $\langle \mathbf{N}, \mathrm{HP} \rangle$ is not strongly successful. But this gives way to a natural

[39] The original incarnation of the Neo-Fregean Program is Wright (1983). For a collection of more recent essays, see Hale and Wright (2001a).

[40] The result, which relies on a one-sorted formulation of HP (see footnote 37), is known as *Frege's Theorem*. It was originally proved by Frege (by making what was later shown to be a non-essential use of Basic Law V), and more recently rediscovered and cleaned-up from contradiction by Crispin Wright. See Frege (1893 and 1903) and Wright (1983).

[41] A different question is whether HP is what underlies our actual knowledge of arithmetic. See Heck (1997a) and Demopoulos (2000).

[42] For a comprehensive survey, see MacBride (forthcoming).

[43] This is only true on a one-sorted reading of HP. But, on a two-sorted version of HP (see footnote 37), HP does not imply the Dedekind-Peano Axioms, so the Neo-Fregean Proposal looses interest.

worry: isn't an *antecedent* justification for the infinity of the universe required for $\langle \mathbf{N}, \mathrm{HP} \rangle$ to deliver a justification for the belief that HP is true? After all, consider the following stipulation:

> The singular term 'Zack' is to be used in such a way that the sentence 'Zack is the man currently in my kitchen' turns out to be true.

Nobody would think that this stipulation can deliver a justification for the belief that Zack is the man currently in my kitchen unless one has an *antecedent* justification for the belief that there is a (unique) man currently in my kitchen. Why should the situation be any different when it comes to $\langle \mathbf{N}, \mathrm{HP} \rangle$? More generally, the following proposal suggests itself:

> *The Straight Proposal*
>
> If a linguistic stipulation is to be strongly successful, the world must cooperate. In particular, the result of existentially Ramsifying the sentence targeted by the stipulation must be true.[44] This is because the target sentence can only be true if its existential Ramsification is true,[45] and strong success can only obtain if the target sentence is rendered true.
>
> In order for the setting forth a stipulation to deliver a justification for believing the stipulation's target sentence, one must have an *antecedent* justification that the world cooperates. In particular, one must have an *antecedent* justification for believing the existential Ramsification of the stipulation's target sentence.
>
> What is true in general is true for $\langle \mathbf{N}, \mathrm{HP} \rangle$. Since the existential Ramsification of HP is equivalent (modulo choice principles) to the infinity of the universe, the setting forth of $\langle \mathbf{N}, \mathrm{HP} \rangle$ can only deliver a justification for believing HP if one has an *antecedent* justification for the infinity of the universe.

The first paragraph of the Straight Proposal should be uncontroversial, even by the lights of Neo-Fregeans. But the final two paragraphs are not, and Neo-Fregeans must find an alternative proposal if their views are to have any plausibility. The problem of finding a plausible Neo-Fregean alternative to the Straight Proposal is the main focus of the present section.[46]

[44]The result of existentially Ramsifying the sentence targeted by the stipulation is the the existential closure of result of uniformly replacing the definienda in the target sentence with variables of the appropriate type.

[45]This makes the simplifying assumption that the target sentence contains no intensional vocabulary, but a similar proposal could be made to accommodate intensional cases.

[46]Much of the material in this section is based on the more detailed discussion in Rayo (2003a).

4.1 Conditional Stipulations

Neo-Fregeans are keen to observe that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is, in a certain sense, *conditional*.[47] More specifically, they argue that $\langle \mathbf{N}, \mathrm{HP} \rangle$ succeeds in assigning a *concept* to '**N**' independently of how matters stand in the world. Facts about the world—in particular, facts about one-one correspondence—enter the picture by determining what *extension* the concept will take.

Whatever the value of this observation in other contexts, it is not helpful as a rebuttal of the Straight Proposal. For, if there are only finitely many objects, then the existential Ramsification of HP is false, and there can be no concept satisfying the constraints imposed by $\langle \mathbf{N}, \mathrm{HP} \rangle$. Thus, insofar as proponents of the Straight Proposal have a warrant for the view that one can only be justified in thinking that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful if one has an antecedent justification for the belief that the universe is infinite, they also have a warrant for the view that one is only justified in believing that $\langle \mathbf{N}, \mathrm{HP} \rangle$ succeeds in creating a concept of the requisite kind if one has an antecedent justification for the belief that the universe is infinite.

4.2 The Abstraction Thesis

Some presentations of the Neo-Fregean Program make heavy use of content-recarving.[48] Specifically, they rely on the view that, as a result of setting forth $\langle \mathbf{N}, \mathrm{HP} \rangle$, $\ulcorner \hat{m}[\mathbf{N}(m, F)] = \hat{m}[\mathbf{N}(m, G)] \urcorner$ comes to have the same content as $\ulcorner F \approx G \urcorner$, carved up in a different way. Call this the *Abstraction Thesis*.

It is important to be clear that the Abstraction Thesis is a *special case* of the view the $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful, since it is a special case of the view that, as a result of setting forth $\langle \mathbf{N}, \mathrm{HP} \rangle$, HP is meaningful and true. Thus, insofar as proponents of the Straight Proposal have a warrant for the view that one can only be justified in thinking that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful if one has an antecedent justification for the belief that the universe is infinite, they also have a warrant for the view that one is only justified in accepting the Abstraction Thesis if one has an antecedent justification for the belief that the universe is infinite.

The Abstraction Thesis is therefore not a good way of countering the Straight Proposal, in the absence of further argumentation. In fact, commitment to the Abstraction Thesis can only complicate the task of offering a Neo-Fregean argument for the view that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful. For however difficult the task might have been when all $\langle \mathbf{N}, \mathrm{HP} \rangle$ is expected to deliver is a material equivalence, matters can only be worse if it is also expected to deliver sameness of content.

It is also worth noting that in section 3.2 we relied on the assumption that HP is true to show that ϕ is an $AB_{\mathcal{F}}$-recarving of $\phi^{\mathcal{F}}$. In particular, we relied on the assumption that HP is true to show that $\ulcorner \hat{m}[\mathbf{N}(m, F)] = \hat{m}[\mathbf{N}(m, G)] \urcorner$ is an $AB_{\mathcal{F}}$-recarving of $\ulcorner F \approx G \urcorner$. So, in the absence of further argumentation,

[47] See, for instance, Hale and Wright (2000), pp. 315-6.

[48] See, for instance, Wright (1997).

the proposal about content-recarving we set forth in section 3 does *not* deliver a justification for believing that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful. It does, however, supply a way of understanding of the Abstraction Thesis according to which the Abstraction Thesis is true on the assumption that $\langle \mathbf{N}, \mathrm{HP} \rangle$ is strongly successful.

Why then is the Abstraction Thesis of any interest? In connection with logicism, there is at least this: if the Abstraction Thesis is true, then a content-preserving paraphrase-function verifies *Language-Logicism* for the special case of numerical identities. Neo-Fregeans sometimes sound as if they believe something further: that if the Abstraction Thesis is true, then there is a sense in which *all* of arithmetic is 'logical'. Perhaps the paraphrase-function $\mathcal{F}$ supplies a way of spelling-out this stronger claim.

4.3 Success by Default

One way of offering an alternative to the Straight Proposal is by defending a *default-epistemology*, according to which it may be sufficient for a belief to be justified that there be no reasons to the contrary.[49] In particular, Neo-Fregeans might attempt to defend something like the following:[50]

> *Success by Default*
>
> In the absence of reasons for thinking that a stipulation fails to meet the adequacy conditions in a certain set S, we are justified in thinking that it is strongly successful.

If *Success by Default* is true, then we needn't have an antecedent justification for the belief that the universe is infinite in order for $\langle \mathbf{N}, \mathrm{HP} \rangle$ to deliver a justification for HP, as proponents of the Straight Proposal would have it. All we need is the absence of reasons for thinking that $\langle \mathbf{N}, \mathrm{HP} \rangle$ fails to meet the conditions in S.

Unfortunately, *Success by Default* is far from being an uncontroversial principle. In order to defend it, Neo-Fregeans must carry out the following two tasks:

1. A specification of the conditions in S must be given.

2. It must be argued that a default epistemology is appropriate for linguistic stipulations.

Much discussion in the Neo-Fregean literature has centered upon subtask 1. The main adequacy condition in S is the requirement that the stipulation in question involve an *abstraction principle*, that is, a principle of the form

$$\forall\alpha\forall\beta\ (\Sigma(\alpha) = \Sigma(\beta) \leftrightarrow R(\alpha, \beta)),$$

[49] I believe this point is due to Crispin Wright. Unfortunately, it does not seem to have been explicitly defended in print.

[50] In a talk entitled 'Implicit Definition and Abstraction' (The University of St Andrews, October 30, 1999), Stewart Shapiro discussed a principle similar to *Success by Default*.

for R an equivalence relation.[51] Unfortunately, this requirement will not do on its own. There are a number of abstraction principles giving rise to stipulations that Neo-Fregeans cannot regard as successful. This is what has come to be known as the *Bad Company Objection*. The simplest example of a problematic abstraction principle is Frege's Basic Law V, which is inconsistent and, hence, unsatisfiable. Neo-Fregeans must therefore add a requirement of consistency to S. But consistency cannot be the end of the matter. Heck (1992), Boolos (1997) and Weir (forthcoming), among others, have set forth an array of consistent but pairwise incompatible abstraction principles. Since *Success by Default* can only be taken to provide support for one or the other of a pair of consistent but incompatible abstraction principles, Neo-Fregeans are forced to postulate additional adequacy conditions. Recent efforts include conservativeness,[52] modesty,[53] stability,[54] and irenicity.[55] But Neo-Fregeans have yet to find a fully satisfactory set of adequacy conditions.

What about task 2? The literature on skepticism offers some general discussion on default epistemologies, mostly in connection with perceptual beliefs.[56] It is argued, for example, that, in the absence of reasons to the contrary, I am now justified in thinking that here is a hand and here is another. Unfortunately, there is no obvious way of extending these conclusions to the realm of linguistic stipulation. And, as far as I know, the adequacy of a default-epistemology for the case of linguistic stipulation has never been defended in print. Unless this omission can be remedied, default-epistemologies are unlikely to pose a serious threat to the Straight Proposal.

4.4 Further Challenges

The Neo-Fregean Program faces at least two further challenges. One concerns second-order logic. Since Neo-Fregeans wish to show that $\langle \mathbf{N}, \mathrm{HP} \rangle$ yields an *a priori* justification for the Dedekind-Peano Axioms, they must argue not only that HP can be known *a priori*, but also that the second-order derivation of the Dedekind-Peano Axioms from HP preserves *a priori* knowledge. Assessing the status of second-order quantification gives rise to difficult questions which

[51] Alternatively, the stipulation might involve a pair of *rules* leading from each side of an abstraction principle to the other.

[52] Roughly, an abstraction principle is conservative if it entails nothing new about objects other than the referents of terms in the principle's left-hand side. See Wright (1997) and Shapiro and Weir (1999).

[53] According to appendix 1 of Wright (1999), "an abstraction [principle] is Modest if its addition to any theory with which it is consistent results in no consequences—whether proof- or model-theoretically established—for the ontology of the combined theory which cannot be justified by reference to its consequences for its own abstracts." Here an abstract is the referent of a term in the left-hand side of the relevant abstraction principle.

[54] An abstraction principle is stable if, for some cardinal κ, it is true at all and only those cardinalities $\geq \kappa$. See Fine (1998), Shapiro and Weir (1999) and Weir (forthcoming).

[55] An abstraction principle is irenic if it is conservative and compatible with any conservative abstraction principle. See Weir (forthcoming).

[56] See, for instance, Wright (2000a), Wright (2000b), Wright (forthcoming) and, for a related proposal, Pryor (2000).

cannot be reviewed here. (See 'Higher-order Logic' and 'Higher-order Logic is not Logic', in this volume.)

Neo-Fregeans also face the *Indeterminacy Challenge*. It arises from the observation that many different assignments of semantic-value to the second-level predicate '**N**' are compatible with the truth of HP. In particular, it is compatible with the truth of HP that the members of any ω-sequence whatsoever serve as referents for the finite numerals (which can easily be defined on the basis of '**N**'). But it would seem to follow, at least at first sight, that the success of $\langle \mathbf{N}, \mathrm{HP} \rangle$ does not by itself suffice to determine whether the referent of '0' is, for instance, Julius Caesar. This is bad news for the Neo-Fregean on the assumption that the following principle obtains:

($*$) An expression cannot be meaningful if there are several different assignments of semantic value, all of which are equally acceptable by the lights of our linguistic practice.[57]

In addressing a related challenge,[58] Hale and Wright have offered an argument to the effect that—contrary to what one might have thought—the success of $\langle \mathbf{N}, \mathrm{HP} \rangle$ does, in fact, suffice to determine that the referent of '0' is not Julius Caesar. It is based on the idea that $\langle \mathbf{N}, \mathrm{HP} \rangle$ establishes one-one correspondence as the 'criterion of identity' for the sortal concept *number*. But suppose it is also conceded that something other than one-one correspondence is the 'criterion of identity' belonging to the sortal concept *person*, and that no object can fall under sortal concepts with different 'criteria of identity'. Then the referent of '0' (which falls under the sortal *number*) must be distinct from Julius Caesar (which falls under the sortal *person*)—or so the argument goes.

Whether or not this sort of argument can be made to work (and whether or not it succeeds in answering the the challenge it was intended to address), it does not suffice to answer the Indeterminacy Challenge. For, as Hale and Wright are aware, their story is compatible with the possibility that the success of $\langle \mathbf{N}, \mathrm{HP} \rangle$ does not determine whether the referent of '0' is the number 7, or any other object falling under the sortal *number*. And, with ($*$) on board, this is enough to give rise to trouble. It is important to be note, however, that that the Indeterminacy Challenge is an instance of a general problem in the philosophy of mathematics,[59] and Neo-Fregeans are free to emulate responses offered by some of their rivals. For example, they can offer resistance to ($*$) by arguing that it threatens the meaningfulness of ordinary English.[60]

[57]Here 'linguistic practice' must be understood broadly enough to include both mental states and environmental factors.

[58]See Hale and Wright (2001b). The related challenge concerns a cluster of interconnected issues which is usually referred to as 'the Caesar-Problem'.

[59]The standard source is, of course, Benacerraf (1965).

[60]See McGee (1993) and Burgess and Rosen (1997) I.A.2.d.

Appendix

Compositionality

We say that a function $*$ is *compositional* with respect to a set S of sentences in a first-order language $\mathcal{L}$ just in case there is:

- a function $\top$;
- a higher-order formula $\xi_\wedge$ containing θ_1 and θ_2 as subformulas and no free occurrences of variables outside θ_1 and θ_2;
- a higher-order formula $\xi_\neg$ containing θ_1 as a subformula and no free occurrences of variables outside θ_1;
- a higher-order formula $\xi_\exists$ containing θ_1 as a subformula and no free occurrences of variables outside θ_1 and θ_2;
- for each n-place atomic predicate P_i^n of $\mathcal{L}$, a formula $\xi_{P_i^n}$ containing free occurrences of all and only the variables $\alpha_1, \ldots, \alpha_n$;

such that the following conditions obtain:[61]

- $\ulcorner \phi \wedge \psi \urcorner^\top = \xi_\wedge[\theta_1/\ulcorner \phi \urcorner^\top][\theta_2/\ulcorner \psi \urcorner^\top]$
- $\ulcorner \neg\phi \urcorner^\top = \xi_\neg[\theta_1/\ulcorner \phi \urcorner^\top]$
- $\ulcorner (\exists x_i)(\phi) \urcorner^\top = \xi_\exists[\alpha/\ulcorner x_i \urcorner^\top][\theta_1/\ulcorner \phi \urcorner^\top]$
- $\ulcorner P_i^n(x_{j_1}, \ldots, x_{j_n}) \urcorner^\top = \xi_{P_i^n}[\alpha_1/\ulcorner x_{j_1} \urcorner^\top] \ldots [\alpha_n/\ulcorner x_{j_n} \urcorner^\top]$
- For any variable $\ulcorner x_i \urcorner$ in $\mathcal{L}$, $\ulcorner x_i \urcorner^\top$ is a variable.
- If ϕ is a sentence of S, then $\phi^\top$ is a sentence.
- For any sentence of S in the range of $*$, $\phi^* = \phi^\top$.

where $\xi[\theta/\phi]$ is the result of replacing every occurrence of θ in ξ by ϕ, and $\xi[\alpha/v]$ is the result of replacing every occurrence of α in ξ by v.

Proposition

Let $\mathcal{L}$ be a first-order language with an intended model $\mathcal{M}$ and suppose that all numerical identities can be interpreted in the theory of $\mathcal{M}$. Let $*$ be a function such that (1) every sentence in $\mathcal{L}$ is in the domain of $*$, (2) there is a higher-order language $\mathcal{L}'$ such that every sentence in the range of $*$ is a sentence of $\mathcal{L}'$, and (3) $*$ is compositional with respect to $\mathcal{L}$. Then, there is no model $\mathcal{S}$ such that $\mathcal{S}$ has a finite domain and, for any sentence ϕ of $\mathcal{L}$, $\models_\mathcal{M} \phi$ if and only if $\models_\mathcal{S} \phi^*$.

[61] For the sake of simplicity, we assume that $\mathcal{L}$ contains no individual constants or function-letters and that all connectives and quantifiers are defined in terms of '$\wedge$', '$\neg$' and '$\exists$', in the usual way.

Proof: If ϕ is a formula of $\mathcal{L}'$ with one free first- or higher-order variable α, say that the *extension* of ϕ relative to $\mathcal{S}$ is the set of semantic-values $s(\alpha)$ for s a variable-assignment satisfying ϕ in $\mathcal{S}$. Since $\mathcal{S}$ has a finite domain, there cannot be an infinite set C such that (a) for some first- or higher-order variable α every member of C is a formula containing only α free, and (b) no two elements of C have the same extension relative to $\mathcal{S}$.

Assume, for *reductio*, that there is a model $\mathcal{S}$ such that $\mathcal{S}$ has a finite domain and such that, for any sentence ϕ of $\mathcal{L}$, $\models_{\mathcal{M}} \phi$ if and only if $\models_{\mathcal{S}} \phi^*$. Since all numerical identities can be interpreted in the theory of $\mathcal{M}$, a formula

$$(\exists x_1)(x_1 = \bar{n} \wedge x_1 = \bar{m})$$

is true in $\mathcal{M}$ whenever $n = m$ and false in $\mathcal{M}$ whenever $n \neq m$. Call this formula $\chi_{n,m}$. By our assumption, $\chi^*_{n,m}$ is true in $\mathcal{S}$ whenever $n = m$, and false in $\mathcal{S}$ whenever $n \neq m$.

Let $n \neq m$. We know that $*$ is compositional with respect to $\mathcal{L}$. Let $\top$ be as in the definition of compositionality. By the definition of compositionality, $\ulcorner x_1 = \bar{n} \urcorner^\top$ and $\ulcorner x_1 = \bar{m} \urcorner^\top$ must contain free occurrences of the variable $\ulcorner x_1 \urcorner^\top$, and no free occurrences of other variables. Moreover, compositionality ensures that $\chi^*_{n,m}$ is the result of substituting $\ulcorner x_1 = \bar{n} \urcorner^\top$ for occurrences of $\ulcorner x_1 = \bar{m} \urcorner^\top$ in $\chi^*_{n,m}$. Since $\chi^*_{m,m}$ is true in $\mathcal{S}$ and $\chi^*_{n,m}$ is false in $\mathcal{S}$, it follows that the extension of $\ulcorner x = \bar{n} \urcorner^\top$ relative to $\mathcal{S}$ is different from the extension of $\ulcorner x_1 = \bar{m} \urcorner^\top$ relative to $\mathcal{S}$.

Consider the set C of formulas $\ulcorner x_1 = \bar{n} \urcorner^\top$, for n a natural number. We know that no two formulas in C have the same extension relative to $\mathcal{S}$. But we had noted above that this is impossible on account of $\mathcal{S}$'s finite domain. $\Box$

Recarving Contents

Let L be an interpreted higher-order language, and allow the variables of L of any given type to be classified according to different 'groups'. Let $\ulcorner \alpha_i^{n,m,k} \urcorner$ be the i-th m-place nth-order variable of the k-th group in L.

Say that a formula '$\alpha_1^{n,m,k} \equiv \alpha_2^{n,m,k}$' is an *identity-predicate* for variables of kind $\ulcorner \alpha_i^{n,m,k} \urcorner$ with respect to a set S just in case (a) '$\alpha_1^{n,m,k} \equiv \alpha_2^{n,m,k}$' is reflexive, transitive and symmetric, and (b) the following condition obtains for any formulas $\ulcorner \phi \urcorner$ and $\ulcorner \psi \urcorner$ in S:

> If $\ulcorner \psi \urcorner$ is the result of replacing every free occurrence of $\ulcorner \alpha_i^{n,m,k} \urcorner$ in $\ulcorner \phi \urcorner$ by $\ulcorner \alpha_j^{n,m,k} \urcorner$ (and perhaps relabelling variables to avoid clashes), then the universal closure of $\ulcorner \alpha_i^{n,m,k} \equiv \alpha_j^{n,m,k} \rightarrow (\phi \leftrightarrow \psi) \urcorner$ is a true sentence of L for any i and j.

Let A and B be sets of formulas of L, and let $\mathcal{I}$ be a compositional,[62] one-one function from A onto B. We shall say that $\mathcal{I}$ *is an identity-isomorphism*

[62] Although the definition of compositionality above is framed in terms of first-order languages, it can easily be generalized so as to encompass higher-order languages.

with respect to A and B just in case there are formulas '$F(\alpha_1^{n,m,k}, \alpha_1^{n',m',k'})$', '$\alpha_1^{n,m,k} \equiv \alpha_2^{n,m,k}$' and '$\alpha_1^{n',m',k'} \equiv' \alpha_2^{n',m',k'}$' of L such that the following conditions obtain:

- '$\alpha_1^{n,m,k} \equiv \alpha_2^{n,m,k}$' is an identity-predicate for variables of kind $\ulcorner \alpha_i^{n,m,k} \urcorner$ with respect to A.
- '$\alpha_1^{n',m',k'} \equiv' \alpha_2^{n',m',k'}$' is an identity-predicate for variables of kind $\ulcorner \alpha_i^{n',m',k'} \urcorner$ with respect to B.
- When $\top$ is as in the definition of compositionality, $\ulcorner \alpha_i^{n,m,k} \urcorner^{\top} = \ulcorner \alpha_i^{n',m',k'} \urcorner$.
- The universal closures of the following are true sentences of L:

  $\forall \alpha_1^{n,m,k} \exists \alpha_1^{n',m',k'} (F(\alpha_1^{n,m,k}, \alpha_1^{n',m',k'}))$
  $\forall \alpha_1^{n',m',k'} \exists \alpha_1^{n,m,k} (F(\alpha_1^{n,m,k}, \alpha_1^{n',m',k'}))$
  $F(\alpha_1^{n,m,k}, \alpha_1^{n',m',k'}) \wedge F(\alpha_2^{n,m,k}, \alpha_2^{n',m',k'}) \rightarrow (\alpha_1^{n,m,k} \equiv \alpha_2^{n,m,k} \leftrightarrow \alpha_1^{n',m',k'} \equiv' \alpha_2^{n',m',k'})$

For $\ulcorner \phi \urcorner \in A$, we say that $\ulcorner \phi \urcorner$ and $\ulcorner \phi \urcorner^{\mathcal{I}}$ are *$AB_{\mathcal{I}}$-equisatisfiable* with respect to variables of kind $\ulcorner \alpha_i^{n,m,k} \urcorner$ if $\mathcal{I}$ is an identity-isomorphism with respect to A and B, and the following additional condition obtains:

- Let $\ulcorner \alpha_1^{n,m,k} \urcorner \ldots \ulcorner \alpha_n^{n,m,k} \urcorner$ be the free variables in $\ulcorner \phi \urcorner$. Then the universal closure of

  $$F(\alpha_1^{n,m,k}, \alpha_1^{n',m',k'}) \wedge \ldots \wedge F(\alpha_n^{n,m,k}, \alpha_n^{n',m',k'}) \rightarrow (\phi \leftrightarrow \phi^{\mathcal{I}})$$

  is a true sentence of L, where $\ulcorner \phi^{\mathcal{I}} \urcorner$ is $\ulcorner \phi \urcorner^{\mathcal{I}}$.

When there is no risk of confusion, we use "$\ulcorner \phi \urcorner$ and $\ulcorner \phi \urcorner^{\mathcal{I}}$ are *equisatisfiable*" as an abbreviation for "$\ulcorner \phi \urcorner$ and $\ulcorner \phi \urcorner^{\mathcal{I}}$ are *$AB_{\mathcal{I}}$-equisatisfiable* with respect to every kind of variables occurring in A".

Finally, we say that $\mathcal{I}$ preserves the *logical networking* of A if $\mathcal{I}$ is an identity-isomorphism with respect to A and B, and $\mathcal{I}$ preserves logical connectives.

References

Ayer, A. J. (1946) *Language Truth and Logic*, Gollancz, London, second edition.

Benacerraf, P. (1965) "What Numbers Could not Be," *The Philosophical Review* 74, 47–73. Reprinted in Benacerraf and Putnam (1983).

Benacerraf, P. (1981) "Frege: The Last Logicist," *Midwest Studies in Philosophy* VI. Reprinted in Demopoulos (1995).

Benacerraf, P., and H. Putnam, eds. (1983) *Philosophy of Mathematics*, Cambridge University Press, Cambridge, second edition.

Boghossian, P., and C. Peacocke, eds. (2000) *New Essays on the A Priori*, Clarendon Press.

Boolos, G. (1984) "To Be is to Be a Value of a Variable (or to be Some Values of Some Variables)," *The Journal of Philosophy* 81, 430–49. Reprinted in Boolos (1998a).

Boolos, G. (1994) "The Advantages of Honest Toil over Theft." In George (1994). Reprinted in Boolos (1998a).

Boolos, G. (1997) "Is Hume's Principle Analytic?" In Richard Heck, *Language, Thought and Logic*, 245-261. Reprinted in Boolos (1998a), 301-314.

Boolos, G. (1998a) *Logic, Logic and Logic*, Harvard, Cambridge, Massachusetts.

Boolos, G. (1998b) "Reply to Parsons' "Sets and Classes"." In Boolos (1998a), 30-36.

Bostock, D. (1979) *Logic and Arithmetic*, Clarendon Press, Oxford.

Burgess, J., and G. Rosen (1997) *A Subject With No Object*, Oxford University Press, New York.

Cartwright, R. (1994) "Speaking of Everything," *Noûs* 28, 1–20.

Dales, H. G., and G. Oliveri, eds. (1998) *Truth in Mathematics*, Oxford University Press, Oxford.

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

Demopoulos, W. (2000) "On the Origin and Status of Our Conception of Number," *Notre Dame Journal of Formal Logic* 41. To appear.

Dummett, M. (1981) *Frege: Philosophy of Language*, Harvard, Cambridge, MA, second edition.

Field, H. (1998) "Which Undecidable Mathematical Sentences have Determinate Truth Values?" In Dales and Oliveri (1998). Reprinted in Field (2001).

Field, H. (2001) *Truth and the Absence of Fact*, Oxford University press, Oxford.

Fine, K. (1998) "The Limits of Abstraction." In Schirn (1998), 503-629.

Frege, G. (1884) *Die Grundlagen der Arithmetik.* English Translation by J.L. Austin, *The Foundations of Arithmetic*, Northwestern University Press, Evanston, IL, 1980.

Frege, G. (1893 and 1903) *Grundgesetze der Arithmetik*. Vol 1 (1893), Vol 2 (1903). English Translation by Montgomery Furth, *The Basic Laws of Arithmetic*, University of California Press, Berkeley and Los Angeles, 1964.

George, A., ed. (1994) *Mathematics and Mind*, Oxford University Press, Oxford.

Glanzberg, M. (typescript) "Quantification and Realism: *Everything* Gets Bigger and Better."

Hale, B. (1997) "Grundlagen §64," *Proceedings of the Aristotelian Society* 97, 243–61. Reprinted in Hale and Wright (2001a), 91-116.

Hale, B., and C. Wright (2000) "Implicit Definition and the A Priori." In Boghossian and Peacocke (2000), 286-319. Reprinted in Hale and Wright (2001a), 117-150.

Hale, B., and C. Wright (2001a) *The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics*, Clarendon Press.

Hale, B., and C. Wright (2001b) "To Bury Caesar. . . ." In Hale and Wright (2001a), 335-396.

Heck, R. (1992) "On the Consistency of Second-Order Contextual Definitions," *Noûs* 26, 491–4.

Heck, R. (1997a) "Finitude and Hume's Principle," *Journal of Philosophical Logic* 26, 589–617.

Heck, R. (1997b) "The Julius Caesar Objection" 273–308. In Heck (1997c).

Heck, R., ed. (1997c) *Language, Thought and Logic*, Clarendon Press, Oxford.

Hodes, H. T. (1984) "Logicism and the Ontological Commitments of Arithmetic," *Journal of Philosophy* 81:3, 123–149.

Hodes, H. T. (1990) "Where do Natural Numbers Come From?" *Synthese* 84, 347–407.

MacBride, F. (forthcoming) "Speaking with Shadows: A Study of Neo-Fregeanism," *British Journal for the Philosophy of Science*.

McGee, V. (1993) "A Semantic Conception of Truth?" *Philosophical Topics* 21, 83–111.

McGee, V. (1997) "How We Learn Mathematical Language," *Philosophical Review* 106, 35–68.

McGee, V. (2000) "'Everything'." In Sher and Tieszen (2000).

Parsons, C. (1974) "Sets and Classes," *Noûs* 8, 1–12. Reprinted in Parsons (1983).

Parsons, C. (1983) *Mathematics in Philosophy*, Cornell University Press, Ithaca, NY.

Pryor, J. (2000) "The Skeptic and the Dogmatist," *Noûs* 34, 517–49.

Rayo, A. (2002) "Word and Objects," *Noûs* 36, 436–464.

Rayo, A. (2003a) "Success by Default?" *Philosophia Mathematica*. To appear.

Rayo, A. (2003b) "When does 'everything' mean *everything*?" *Analysis* 63. To appear.

Rayo, A. (forthcoming) "Frege's Unofficial Arithmetic," *The Journal of Symbolic Logic*.

Rayo, A., and G. Uzquiano (1999) "Toward a Theory of Second-Order Consequence," *The Notre Dame Journal of Formal Logic* 40, 315–325.

Rayo, A., and T. Williamson (forthcoming) "A Completeness Theorem for Unrestricted First-Order Languages."

Schirn, M., ed. (1998) *Philosophy of Mathematics Today*, Clarendon Press, Oxford.

Shapiro, S. (1991) *Foundations Without Foundationalism: A Case for Second-Order Logic*, Clarendon Press, Oxford.

Shapiro, S., and A. Weir (1999) "New V, ZF and Abstraction," *Philosophia Mathematica* 7:3, 293–321.

Sher, G., and R. Tieszen, eds. (2000) *Between Logic and Intuition*, Cambridge University Press, New York and Cambridge.

Uzquiano, G. (2002) "Categoricity Theorems and Conceptions of Set," *The Journal of Philosophical Logic* 31:2, 181–196.

Weir, A. (forthcoming) "Neo-Fregeanism: An Embarrassment of Riches," *The Notre Dame Journal of Formal Logic*.

Whitehead, A., and B. Russell (1910-13) *Principia Mathematica*, Cambridge University Press, Cambridge, 2nd edition. 3 vols.

Williamson, T. (1999) "Existence and Contingency," *Proceedings of the Aristotelian Society, Supplementary Volume* 73, 181–203. Reprinted with printer's errors corrected in *Proceedings of the Aristotelian Society* 100 (2000), 321–343 (117–139 in unbounded edition).

Williamson, T. (typescript) "Everything."

Wright, C. (1983) *Frege's Conception of Numbers as Objects*, Aberdeen University Press, Aberdeen.

Wright, C. (1997) "The Philosophical Significance of Frege's Theorem." In Heck (1997c), 201-244.

Wright, C. (1999) "Is Hume's Principle Analytic," *Notre Dame Journal of Formal Logic* 40, 6–30. Reprinted in Hale and Wright (2001a), 307-332.

Wright, C. (2000a) "Cogency and Question-Begging: Some Reflections on McKinsey's Paradox and Putnam's Proof," *Philosophical Issues* 10, 140–63.

Wright, C. (2000b) "replies," *Philosophical Issues* 10, 201–19.

Wright, C. (forthcoming) "Wittgensteinian Certainties." to appear in McManus D. (ed.) *Wittgenstein and Scepticism.*

Yablo, S. (2002) "Abstract Objects: A Case Study," *Philosophical Issues*. Originally appeared in Giarrette, D. (ed.) *Individuals, Essence and Identity.*

Zermelo, E. (1930) "Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre," *Fundamenta Mathematicae* 16, 29–47.