Frege's Context Principle and Reference to Natural Numbers*

Øystein Linnebo
University of Bristol

Draft of 30 June 2006

Abstract

Frege proposed that his Context Principle—which says that a word has meaning only in the context of a proposition—can be used to explain reference, both in general and to mathematical objects in particular. I develop a version of this proposal and outline answers to some important challenges that the resulting account of reference faces. Then I show how this account can be applied to arithmetic to yield an explanation of our reference to the natural numbers and of their metaphysical status.

At *Grundlagen* [6] §62 Frege raises a question that has dominated much of recent philosophy of mathematics: "How, then, are the numbers to be given to us, if we cannot have any ideas or intuitions of them?" The problem is of course that numbers, unlike tables and chairs, cannot be perceived; nor can they be observed with the help of modern technology, as electrons and DNA molecules can. How can we then refer to numbers and other kinds of abstract objects, let alone gain knowledge of them?

It is important to get clear on the nature of this question. What the question calls for is not an account of *what objects* various numerals (or their mental counterparts) refer to. Giving such an account is easy; for instance, '7' and 'VII' refer to 7. Rather, what the question calls for is an account of what facts about reference *consist in.* For instance, we would like to know how it comes about that '7' and 'VII' refer to 7. This question is a perfectly reasonable one. For the fact that a term or representation manages to refer to an object external to itself can hardly be a primitive fact but must have some explanation.

Frege's next sentence makes a proposal about how this question can be addressed. "Since it is only in the context of a sentence that words have any meaning, our problem becomes this: To define the sense of a sentence in which a number word occurs." The doctrine that

*Thanks to Matti Eklund, Anthony Everett, Agustín Rayo, Stewart Shapiro, and Timothy Williamson for discussion and comments on earlier versions of this paper. Thanks also to audiences at Uppsala University and the University of Oxford, where earlier versions of this paper were presented.

words have meaning only in the context of a sentence has become known as the *Context Principle*.[1] What Frege proposes is that the Context Principle has an essential role to play in the explanation of reference, both in general and to numbers and other abstract objects in particular. I will refer to this as *Frege's Proposal*. The idea is to translate the problem of explaining how a singular term comes to refer into the problem of explaining how certain complete sentences involving this term come to be meaningful.

In this paper I develop an account of reference based on Frege's Proposal (Section 1) and outline answers to some important challenges that this account faces (Section 2). Then I show how this account can be applied to arithmetic to yield an explanation of our reference to the natural numbers (Section 3) and of their metaphysical status (Section 4). Given the daunting magnitude and difficulty of the questions I discuss, all I can reasonably hope to do in this paper is to outline a research programme and to begin exploring its strengths and possible weaknesses. A more complete treatment of the questions I discuss would no doubt require an entire book. I should also mention that my present goal is systematic, not exegetical. Although I believe Frege anticipated many aspects of the account that I develop, I do not claim that he would have agreed with all of it.

1 Towards a Fregean account of reference

Assume we want to explain how singular terms of a certain kind come to refer. According to Frege's Proposal, we can do this by explaining how complete sentences involving this kind of singular terms come to be meaningful. Before we attempt to provide any such explanations, it will be useful to simplify the problem somewhat.

The first simplification is to focus on thought rather than on language. Our modified *explanandum* is then what is involved in someone's capacity for singular reference to various sorts of object.[2] The proposal is that an adequate explanation of this can be given by explaining what is involved in the person's capacity for understanding complete thoughts concerning objects of the sort in question.[3] This simplification will allow us to concentrate on an indi-

[1] See also *ibid.* pp. x, 71, and 116. I have changed the translation of 'Satz' from 'proposition' to 'sentence'. This is reasonable, given that Frege talks about *words* occurring in a *Satz*.

[2] In the terminology of [4], my goal is to explain what our understanding of the relevant kind of "fundamental Ideas" consists in.

[3] Strictly speaking, I here collapse two steps. The first step is Frege's suggestion that questions concerning singular reference be addressed in terms of analogous questions concerning complete thoughts. In particular, in virtue of what does a physical state of an agent have a particular thought as its content? The second step is to approach this question about thoughts in terms of the notion of understanding. Doing so is quite natural; for in order to stand in some propositional attitude to a thought, one presumably needs to *understand* that thought.

vidual person rather than on a whole language community. This is a huge simplification. For instance, if reference involves a notion of Fregean sense, then this sense will now be allowed to vary with each individual act of reference. In contrast, if a notion of sense is to be attached to an expression of a public language, then this sense will have to be shared by every competent speaker of this language. (For the purposes of this paper I will not have anything to say about how linguistic expressions come to refer.)

The second simplification is to focus on *canonical* cases of singular reference. These are certain maximally direct ways of referring to objects, where the referent is "directly present" to the thinker. For instance, referring to a person whom I see immediately in front of me is canonical, whereas referring to Napoleon is not. More examples of canonical reference will be presented shortly.[4]

The third simplification is to begin by explaining someone's understanding of identity statements before attempting to explain his understanding of thoughts more generally. This strategy is adopted by Frege himself in *Grundlagen*.[5] The rationale is that, before one can understand what it means for an object to possess properties and stand in relations, one needs to know how to distinguish the object from other objects and how to re-identify it when presented with it in alternative ways.[6]

With these three modifications, Frege's Proposal becomes the following: We can explain what is involved in someone's capacity for canonical singular reference to objects of a certain kind by explaining what is involved in his or her capacity for understanding identities concerning such objects. This would translate the problem of explaining our capacity for singular reference into the related but different problem of explaining our capacity for understanding identity statements.

Frege suggests an ingenious way in which his proposal can be carried out. The core idea is that canonical reference has a rich and systematic structure. Firstly, objects are always presented to us only via some of their parts or aspects. And secondly, we have a grasp of how two such parts or aspects must be related for them to pick out the same object. Here are some examples.[7]

---

[4]Following Michael Dummett and Gareth Evans I believe non-canonical reference must be explained in terms of someone's ability to recognize the referent when presented with it in a canonical way. See [2], pp. 231-239 and [4], pp. 109-112. (In a more complete treatment, I would attempt to give a more explicit account of non-canonical reference.)

[5]Frege unfortunately abandons this strategy in *Grundgesetze*. For an analysis, see my [9].

[6]Cf. Evans, *op. cit.*, who explains fundamental Ideas in terms of "fundamental grounds of difference."

[7]In a more complete treatment, each example would of course have to be developed in greater detail and defended against objections. My present goal is merely to sketch some promising examples in order to illustrate how the Fregean framework functions.

(a) *Physical bodies.* A physical body is most directly presented in perception, where we causally interact with one or more of its spatiotemporal parts. Two such parts determine the same physical body just in case they are connected through a continuous stretch of solid stuff, all of which belongs to a common unit of motion.[8]

(b) *Directions.* A direction is most directly presented by means of a line (or some other directed object) that has the direction in question. Two lines determine the same direction just in case they are parallel.

(c) *Shapes.* This case is analogous to that of directions: Shapes are most directly presented by things or figures that have the shape in question. Two such things or figures determine the same shape just in case they are congruent.

(d) *Syntactic types.* Syntactic types are most directly presented by means of their tokens. Two tokens determine the same type just in case they count (according to the relevant standards) as instantiating the same type.

(e) *Natural numbers.* A natural number is most directly presented by means of some member of a sequence of numerals. Two numerals determine the same number just in case they occupy analogous positions in their respective sequences.[9]

These examples suggest that canonical cases of singular reference are always based on two elements. First, there is an intermediary entity in terms of which the referent is most immediately presented. Let's call this the *presentation*. Second, there is a relation which specifies the condition under which two presentations determine the same referent. Let's refer to this as the *unity relation*. Finally, let's call an ordered pair $\langle u, \approx \rangle$ consisting of a presentation u and a unity relation $\approx$ applicable to this presentation a *referential attempt*. Frege's proposal is then that canonical reference is based on referential attempts. Does this proposal yield an adequate explanation of what someone's understanding of singular reference consists in?

A *formal* adequacy condition is obviously that the account be non-circular. It is easily seen that the form of our proposal allows it to be non-circular. Consider for instance the case

[8] I elaborate on this view and defend it against some natural objections in my [12].

[9] I elaborate on this view in Section 3 below; related ideas are found in [14]. This view of the natural numbers as finite *ordinals* contrasts with the logicist view that the natural numbers are finite *cardinals*, individuated by Hume's Principle (which says that two numbers are identical just in case the concepts whose numbers they are are equinumerous). However, both views are compatible with the Fregean account of reference. It is thus largely an empirical question which view best describes human thought about the natural numbers. See Section 3.

of directions. What is proposed is that someone's understanding of an identity statement concerning directions can be explained in terms of his being suitably related to lines (in terms of which directions are presented) and having a suitable grasp of parallelism (which is the unity relation). In this case there is no threat of circularity, as we can explain what it is for someone to be suitably related to lines and to have a suitable grasp of parallelism without presupposing any prior ability to explain reference to directions. Next, we observe that there is nothing in this example that is peculiar to the case of directions. Our proposal is to explain someone's understanding of an identity statement in terms of his being suitably related to the relevant presentations and having a suitable grasp of the relevant unity relation. This explanation will of course have to include an account of what it is for a person to be suitably related to these presentations and to have a suitable grasp of this unity relation. But there is no general reason why this account should presuppose what we are trying to explain, namely reference to the sort of objects that are determined by these presentations and this unity relation.[10]

The *material* adequacy condition is that the account should capture what someone's capacity for singular reference consists in. My argument that this adequacy condition is satisfied is based on two claims. First I claim that my account explains what the subject's understanding of identity statements involving the referent consist in. Consider a representation **a** purporting to make singular reference to some object. According to my account, this representation is associated with some referential attempt $\langle u, \approx \rangle$, which specifies how the referent is presented and when two such presentations determine the same referent. By operating with this referential attempt, the subject will be able to understand any thought of the form $\ulcorner \mathbf{a} = \mathbf{b} \urcorner$, where **b** is any other representation purporting to make singular reference to an object of the kind in question. For according to my account, **b** too must be associated with some referential attempt, say $\langle v, \approx \rangle$. Moreover, we are assuming that the subject operates correctly with these representations, namely in accordance with the following principle for the identity of their semantic values:

$$\text{(SV)} \qquad [\![\mathbf{a}]\!] = [\![\mathbf{b}]\!] \leftrightarrow u \approx v$$

This means that the subject has an ability to track the referent of **a** and to distinguish it from other objects of the same sort.

[10]This is of course not to say that there cannot be *particular* cases where such an illicit presupposition exists. In fact, in Section 2.2 I will suggest that some problems encountered by Frege's Proposal are caused by the use of presentations and unity relations an adequate grasp of which *would* presuppose an ability to refer to the entities in question, thus making the account viciously circular.

Next I claim that this competence is naturally described as knowing what object the representation **a** refers to. Consider for instance the case of physical bodies. Assume someone is digging in the garden, hits upon something hard with her shovel, and as a result forms the thought: This body is large. Later she hits upon something hard again, one meter away from the first encounter, and as a result forms the thought: This body is identical to that body. Finally, our subject appreciates that this identity statement is true just in case the two chunks of solid stuff that she has hit upon are spatiotemporally connected in the suitable way. It is extremely plausible to describe this capacity as a capacity to refer to physical bodies. For instance, if a robot was equipped with perception-like mechanisms and programmed so as to operate with the appropriate unity relation, it would make sense to ascribe to the robot a basic capacity for referring to physical bodies.

More generally, the unity relation $\approx$ implicitly defines a (partial) function f that maps a presentation u to the referent, if any, that u picks out. This is encapsulated in what I will call *principles of individuation*:[11]

$$f(u) = f(v) \leftrightarrow u \approx v \tag{PI}$$

Of course, when formulating principles of individuation, we philosophers make use of our own ability to refer to objects of the kind in question. But this is perfectly permissible. We are allowed to presuppose *that* we can refer to objects of the kind in question. What we are not allowed to presuppose is an explanation of what this ability *consists in*. But no such presupposition is made.

Principles of individuation are obviously closely related to what are known as *abstraction principles*, that is, to principles of the form

$$\Sigma(\alpha) = \Sigma(\beta) \leftrightarrow \alpha \sim \beta \tag{$*$}$$

where Σ is a term-forming operator, α and β are (either first- or second-order) variables, and where $\sim$ denotes an equivalence relation on the entities that α and β range over. However, the notion of an abstraction principle is a purely logical one, which applies to any principle of the form just described. By contrast, the notion of a principle of individuation is a philosophical one, which applies to a principle of the requisite form just in case it can serve in a Fregean account of reference.

[11] Principles of the form (PI) are sometimes called *two-level criteria of identity.* For two approaches based on such principles, see [16] and [19], Chapter 9.

This principal difference brings with it some other differences as well. For instance, the operator Σ in an abstraction principle is required to be defined on all values of the variables α and β, whereas the function f in a principle of individuation is allowed to be partial. We allow this because some presentation fail to determine referents; there are for instance spatiotemporal chunks that fail to determine unique bodies. This has consequences for the unity relation $\approx$ as well. Recall that the unity relation gives the condition under which two presentations determine the same referent. This means that any unity relation has to be symmetric and transitive. However, the unity relation will be reflexive only on those presentations that succeed in determining referents. For a referential attempt $\langle u, \approx \rangle$ succeeds in determining a referent just in case the presentation u bears the unity relation $\approx$ to itself. It follows that the domain of the partial function f is identical to the field of the unity relation $\approx$.[12] (In Section 2.2 we will encounter another difference as well, namely that the Non-Circularity Constraint puts serious restrictions on what abstraction principles can serve as principles of individuation.)

2 Challenges to the Fregean account of reference

In the previous section I described Frege's Proposal and outlined an attempt to carry out this Proposal. Although I argued that this attempt looks quite promising, a large number of challenging questions remain. I will now discuss four such questions, which are all concerned with very general features of Frege's Proposal. These are hard questions, each deserving a paper-length response of its own. All I can do here is to outline some responses and thus make it plausible that my account has the resources needed to address these worries.

2.1 Basic reference

When a subject makes a singular reference based on a referential attempt $\langle u, \approx \rangle$, what is the nature of his relation to the presentation u? Here are three possible answers.

The first answer is that the subject's relation to the presentation u is one of full-fledged singular reference. My account of reference can then be applied again, yielding the conclusion that reference to u has to be based on some further presentation and an appropriate unity relation. Such iterated applications of my account of reference are clearly possible. For instance, in example (e) of the previous section I claimed that natural numbers are picked

[12] Recall that the *field* of a relation R is the set of objects which R relates. Thus, when R is dyadic, its field is the set $\{x \mid \exists y(Rxy \vee Ryx)\}$. Note that it is a theorem of first-order logic that a symmetric and transitive relation is reflexive on all objects in its field.

out by means of numeral types, which in turn are picked out by means of numeral tokens. However, on pain of a vicious regress, this cannot be a complete answer; in particular, it cannot explain how reference comes about in the first place. So at least in some cases, the subject's relation to a presentation u must be something less than full-fledged singular reference. I will say that a case of full-fledged singular reference is *basic* when it doesn't depend on any other cases of such reference. What I have argued is thus that the first answer is fine as far as it goes but that it cannot account for basic reference.

The second answer holds (like the first) that the subject's relation to the presentation u is a referential one, but insists (unlike the first) that the presentation u need not be fully individuated. Since the Fregean account of reference applies only to objects that are fully individuated, this allows us to avoid a vicious regress. For when a presentation u isn't fully individuated, the Fregean account cannot be applied to u. Instead the presentation u will contribute to a instance of basic reference.

However, it is far from clear how the not-fully-individuated presentations are to be understood. Let's focus on the clearest example of basic reference, namely reference to physical bodies. We would like to know how the required talk of spatiotemporal parts and spatiotemporal continuity can be understood in a pre-individuative manner. Probably the most developed reponse is due to Michael Dummett, who uses as his paradigm what P.F. Strawson calls "feature-placing" sentences, such as 'It is wet here' or 'It is hot there'.[13] Such sentences involve no singular reference at all: The pronoun 'it' is merely a formal subject, not a semantic one. The same goes, according to Dummett, for the claim that 'this is continuous with that' (accompanied by two pointing gestures). This claim doesn't involve any form of singular reference to determinate parts but only feature-placing. However, it is not clear that this response delivers what we need. To begin with, this would mean that the variables u and v in the principles of individuation cannot always be ordinary first-order variables; for such variables range over ordinary (and thus "fully individuated") objects. But more seriously, it is unclear how we are to understand talk about and quantification over some class of presentations before it has been determined what it is for two such presentations to be identical or distinct.

The third answer to our question is that a subject's relation to a presentation isn't a referential one at all but something more primitive. I believe this answer provides the most promising account of basic reference. By denying that the relation between a subject and a presentation need be a referential one, this answer allows us to avoid a vicious regress. And it

[13] See [2], p. 217 and [3], pp. 162f.

does so in a way that avoids any vague talk about not-fully-individuated presentations. But if the relation between a subject and a presentation isn't a referential one, what then is it? In the case of reference to physical bodies, the subject's relation to the spatiotemporal parts that serve as presentation can plausibly be taken to be a purely causal one: It is the relation that holds between the subject's perceptual system and the sum-totals of particle-instants with which the subject causally interacts in the appropriate perceptual way. In cases of reference that rely on concepts as presentations, the relation between subject and presentation is that of concept possession. Granted, it is far from clear how concept possession is to be explained; and I have no account on offer here. Even so, it is reasonably clear that no plausible account of concept possession will identify the *possession* of a concept with *reference* to that concept. This provides another example of how the relation between a subject and a presentation need not be a referential one.

2.2 The Bad Company Problem

Another problem faced by my Fregean account of reference is the need for a principled demarcation of the principles of individuation that are acceptable from those that are not. The most serious aspect of this problem has to do with *the consistency* of the underlying abstraction principles. For although the abstraction principles considered in Section 1 are consistent, it is well known that many others are not. The most famous example of this is Frege's Basic Law V, which says that the extensions of two concepts are identical just in case the two concepts are coextensive:

$$\text{(V)} \qquad \hat{x}.Fx = \hat{x}.Gx \leftrightarrow \forall x(Fx \leftrightarrow Gx)$$

For in second-order logic, (V) allows us to derive Russell's paradox.[14] The problem of giving an informative characterization of the acceptable abstraction principles is known as *the Bad Company Problem.*

I am hopeful that this problem can be solved by carefully heeding the requirement that our Fregean account of reference be non-circular. Recall that this account seeks to translate the problem of explaining how reference comes about to the related but different problem of explaining what someone's understanding of identity statements concerning such objects consists in. If this translation is to constitute progress, the desired explanation of our capacity for understanding identity statements must obviously not presuppose any prior ability to

[14]The derivation of Russell's paradox requires Σ^1_1-comprehension. [18] proves that the second-order theory with Δ^1_1-comprehension and (V) as its sole non-logical axiom is consistent.

explain our capacity for making singular reference to objects of the kind in question. I will refer to this as *the Non-Circularity Constraint.* Note that this constraint does not prevent us from presupposing *that* we are capable of making the relevant sort of singular reference. All it disallows is presupposing an ability to explain *what this capacity consists in.*[15]

The Non-Circularity Constraint requires that, when someone makes a referential attempt $\langle u, \approx \rangle$, it be possible for them to stand in an appropriate relation to the presentation u and have an appropriate grasp of the unity relation $\approx$ without already being capable of referring to the kind of objects reference to which we are attempting to explain. This means that the presentation u and the unity relation $\approx$ cannot involve or presuppose such objects as figure on the left-hand side of the principle of individuation.

Sets provide a nice example of how the Non-Circularity Constraint works. I believe a set is presented by means of the plurality of its elements, and that two pluralities determine the same set just in case they encompass the same objects. (The pluralities in question can be represented by means of plural variables, as suggested by George Boolos.[16]) We may also need to add that only pluralities that satisfy some condition of "limitation of size" succeed in determining sets. Let's write $\textsc{Form}(uu, x)$ for the claim that the objects uu form a set x; that is, that there is a set x whose elements are precisely the objects uu. My claim is then that sets are individuated as follows

$$\text{(Id-Sets)} \qquad \textsc{Form}(uu, x) \wedge \textsc{Form}(vv, y) \rightarrow (x = y \leftrightarrow uu \equiv vv)$$

where '$uu \equiv vv$' is a formalization of the claim that the pluralities uu and vv encompass precisely the same objects.[17] Let's now apply the Non-Circularity Constraint. This constraint requires that we be able to refer to the individual objects that make up a plurality *before* we can refer to the set that this plurality forms. This means that a set cannot contain itself as an element. More generally, the constraint can be seen to give rise to the set-theoretic axiom of Foundation.

Extensions, on the other hand, are individuated by Frege's Basic Law V: They are presented by means of concepts and are identical just in case the two presenting concepts are coextensive. But unless (V) is restricted it in some way it will lead straight to paradox, as Russell discovered. My hypothesis is that the problem with the unrestricted version of (V) is that it violates the Non-Circularity Constraint: that its right-hand side somehow presupposes

[15] Recall that this distinction played an important role in my defense of (PI) towards the end of Section 1.
[16] See [1]; for an introduction, see [10].
[17] For those familiar with plural logic, this claim can be formalized as $\forall z(z \prec uu \leftrightarrow z \prec vv)$.

an understanding of what we are attempting to explain on the left-hand side, namely what it is for two extensions to be identical.

One way of developing this suggestion is as follows. Say that an abstraction principle is *predicative* when its unity relation doesn't quantify over the kinds of entities to which its left-hand side purports to refer, and that it is *impredicative* otherwise. For instance, the abstraction principle that specifies how directions are individuated from lines,

$$d(l_1) = d(l_2) \leftrightarrow l_1 \parallel l_2 \tag{D}$$

is predicative, as there is no reference to, or quantification over, directions on its right-hand side. By contrast, (HP) and (V) are both impredicative, as their right-hand sides quantify over precisely the kinds of object that their left-hand sides attempt to individuate. Could it be that the unity relations involved in impredicative abstraction principles are illegitimate because they presuppose an ability to individuate the objects in question, thus violating the Non-Circularity Constraint?

The following argument appears to suggest that this is indeed so. Let (∗) be an impredicative abstraction principle. The impredicativity means that the right-hand side of (∗) quantifies over the very sort of objects, say Fs, reference to which we are attempting to explain. Now it appears that in order to understand a quantified formula, one must be able to understand an arbitrary instance of this formula. If this is right, it follows that one cannot understand the right-hand side of (∗) unless one is already able to understand instances, with respect to Fs, of the formulas to which its quantifiers attach. But in order to understand such instances, one needs to be capable of making singular reference to Fs. This means that use of an impredicative abstraction principle as a principle of individuation is circular: it presupposes precisely what it attempts to explain, namely an understanding of singular reference to the relevant sort of objects.[18]

So if this argument were sound, only predicative abstraction principles would be permitted as principles of individuation. However, a crucial step in this argument fails: It is *not* the case that, in order to understand a quantified formula, one needs to be capable of understanding an arbitrary instance. We *can* understand universal generalizations without being able to understand an arbitrary instance. For instance, we can understand (and even appreciate the truth of) the claim that all objects are self-identical, although there are no doubt many kinds of object to which we will never be capable of referring. What does the work in such cases is

[18] A similar argument is developed in [5], Section II.5, whereas the opposite view is defended in [21] and [22].

that we have an absolutely general understanding of what it is for the relevant condition to hold of an object. This is what enables us to tell that *any object whatsoever* is self-identical, including objects of which we will never gain more substantive knowledge.

We therefore need a better analysis of what a quantified formula presupposes. I will here focus on the presuppositions carried by the open formulas (or *conditions* as I will also call them) that define concepts and extensions. In particular, what do such conditions presuppose about the identities of the objects on which they are defined? Since what a condition does is distinguish between objects—those of which it holds and those of which it doesn't—this notion of presupposition should be spelled out in terms of what distinctions the condition makes. Now, a condition can only presuppose an object if it is able to distinguish this object from other known objects. I therefore propose that we analyze what it is for a condition to presuppose only entities that are already individuated in terms of the condition's not distinguishing between entities not yet individuated. This proposal can be made mathematically precise as follows. Consider permutations π that fix all objects already individuated and that respect all relations already individuated in the sense that for each such relation R we have

$$\forall x_0 \ldots \forall x_n (Rx_0 \ldots x_n \rightarrow R\pi x_0 \ldots \pi x_n).$$

My proposal is then that a condition $\phi(u)$ presupposes only entities that are already individuated just in case $\phi(u)$ is invariant under all such permutations. The "Russell condition" $\exists F(u = \hat{x}.Fx \wedge \neg Fu)$ provides an interesting example. This condition is easily seen to violate the present analysis of the Non-Circularity Constraint; for this condition is sensitive to the characteristics of the object assigned to the variable u. So this condition cannot be allowed to define a concept or an extension.

When a condition does satisfy the Non-Circularity Constraint, there is no obvious philosophical reason why it should not define an extension. For this extension has been individuated in a non-circular way. Nor is there any mathematical reason why such a formula should not define a extension.[19] To see this, begin by observing that such a condition defines a *concept*. Next I claim that any concept individuated in accordance with the Non-Circularity Constraint can be assigned an extension. For each such concept can be represented by means of one of the objects not yet individuated. Since these concepts don't distinguish between objects not yet individuated, it doesn't matter which representative we choose. So if we assume that there are as many objects not yet individuated as there are sets (an assumption that can eas-

[19] What follows is an intuitive presentation of ideas developed in more detail in Sections 7 and 8 of my [13].

ily be modeled within set theory), there will be enough objects to represent all the concepts definable by conditions expressible in any reasonable language. And we can carry out this process as many times as there are sets. I therefore conclude that when the individuation of concepts is subjected to the Non-Circularity Constraint, the resulting concepts do indeed have extensions in accordance with Basic Law V.

2.3 Compositionality

I now turn to a worry about the compatibility of my Fregean account of reference with the principle of compositionality. This worry is based on the following observation. According to the principle of compositionality, the meaning of anything complex is to be explained in terms of the meanings of its constituent parts. So here the order of explanation goes from what is complex to what is simpler. But according to my account of reference, the referentiality of a singular representation is partially explained in terms of the meaningfulness of identities involving this representation. So here the order of explanation goes in the opposite direction: from what is absolutely simple to what is more complex. Given these opposite orders of explanation, one may think my account conflicts with the principle of compositionality.

One way *not* to resolve this apparent conflict would be by claiming that, whereas my account is only concerned with thought, the principle of compositionality governs only linguistic meaning. This response is unacceptable because an analogous principle of compositionality applies to the contents of thoughts.

However, there is another respect in which the principle of compositionality and my Fregean account of reference do have completely different concerns. To see this, we need to distinguish between what I will call *semantics* and *meta-semantics*.[20] Semantics standardly takes the form of a theory of *semantic values*, where the semantic value $[\![\mathbf{E}]\!]$ of an expression $\mathbf{E}$ is the contribution that this expression makes to the truth-values of sentences in which it occurs.[21] Following Frege, the semantic value of a sentence is often taken to be just its truth-value, and the semantic value of a proper name, its referent. If so, it follows that the semantic value of a one-place predicate must be a function from objects to truth-values. Following Frege again, it is argued that semantic values are subject to a *principle of compositionality*, according to which the semantic value of a complex expression is determined as a function of the semantic values of its individual sub-expressions. For instance, the semantic value of an atomic sentence $\mathbf{P}(\mathbf{a_1}, \ldots, \mathbf{a_n})$ is functionally determined as $[\![\mathbf{P}]\!]([\![\mathbf{a_1}]\!], \ldots, [\![\mathbf{a_n}]\!])$.

[20]My distinction between semantics and meta-semantics is thus the same as Stalnaker's distinction between "descriptive" and "foundational" semantics. See e.g. [17].

[21]I will use boldface for all meta-linguistic variables.

Meta-semantics, on the other hand, is concerned with what is involved in an expression's having the various semantic properties that it happens to have, such as its semantic structure and its semantic value. The relation between semantics and meta-semantics can be compared with that between economics and what we may call *meta-economics*. Economics is concerned with the laws governing money; for instance, that an excessive supply of money leads to inflation. Meta-economics, on the other hand, is concerned with what is involved in various objects' having monetary value; for instance, what makes it the case that a piece of printed paper can be worth 100 euros. Clearly, such facts cannot be primitive but must have an explanation.

The meta-semantic questions that we are currently interested in concern singular terms and representations. How does the relation of reference come about? What makes it the case that a "dead" syntactic object—some ink marks on paper or neural configurations in the brain—"reach out" to some referent with which the singular term bears no intrinsic connection? Just like the fact about the value of a banknote, this semantic fact cannot be primitive but must have an explanation.

I would instead like to suggest that our distinction between semantics and meta-semantics provides the key to resolving the apparent conflict between my Fregean account of reference and the principle of compositionality. The principle of compositionality is concerned with the assignment of semantic values to complex expressions and thus belongs to semantics. My Fregean account of reference, on the other hand, is concerned with what is involved in an expression's having the various semantic properties it happens to have and thus belongs to meta-semantics. Since the principle of compositionality and our Fregean account of reference have completely different concerns, there is no conflict.

However, the most popular response among philosophers who seek to use Frege's Context Principle to explain reference has been to concede that the apparent conflict is genuine and therefore to argue that the principle of compositionality has to be rejected or at least weakened.[22] But rejecting or weakening the principle of compositionality is obviously a steep price to pay. Why, then, have so many philosophers found this response inevitable? I believe the answer has to do with a dangerous ambiguity in the wording of Frege's proposal. "Since it is only in the context of a sentence that words have any meaning," Frege writes, "our problem becomes this: *To define the sense of a sentence in which a number word occurs.*"[23] This is ambiguous between a semantic and a meta-semantic reading. On the semantic reading, our task is to *specify* the meaning or sense of identity statements in which number words occur.

[22] See [7] and [20].

[23] *Grundlagen* §62; my italics.

But on the meta-semantic reading, our task is to *explain what makes it the case that such identity statements have the meanings that they happen to have.*

Frege's proposal has traditionally been interpreted along the lines of the semantic reading. It is then natural to assume that what Frege proposes is that the meaning of problematic identity statements be given by a "reductive" truth-condition

(T-Red) $$\ulcorner \mathbf{a} = \mathbf{b} \urcorner \text{ is true iff } u \approx v$$

where **a** and **b** are representations associated with referential attempts $\langle u, \approx \rangle$ and $\langle v, \approx \rangle$ respectively. On this reading there will indeed be a conflict with the principle of compositionality. For according to this principle, the semantic value of an atomic sentence $\mathbf{P}(\mathbf{a_1}, \ldots, \mathbf{a_n})$ is functionally determined as $[\![\mathbf{P}]\!]([\![\mathbf{a_1}]\!], \ldots, [\![\mathbf{a_n}]\!])$. Applied to the identity $\ulcorner \mathbf{a} = \mathbf{b} \urcorner$, this yields something different from (T-Red), namely the completely trivial truth-condition

(T-Triv) $$\ulcorner \mathbf{a} = \mathbf{b} \urcorner \text{ is true iff } [\![\mathbf{a}]\!] = [\![\mathbf{b}]\!].$$

And any further or alternative semantic analysis is out of the question, given that the terms **a** and **b** are supposed to be semantically simple. Moreover, the truth-condition (T-Triv) will be of absolutely no use in the project of explaining some problematic form of reference. For the right-hand side of (T-Triv)—unlike that of (T-Red)—involves precisely the sort of reference that we are attempting to explain.

Faced with this choice between the reductive truth-condition (T-Red), which allows the explanatory project to progress, and the trivial one (T-Triv), on which the explanation cannot even get started, it is of course tempting to insist that it is the former that gives the meaning of the identity statement, and that if this conflicts with the compositionality of meaning, then so much the worse for this principle of compositionality. This appears to have been Frege's view in *Grundlagen*, where he talks about the right-hand side being a "recarving" of the meaning of the left-hand side.[24] This "recarving thesis" is explicitly endorsed by prominent contemporary defenders of Fregean ideas about reference, such as Bob Hale and Crispin Wright.[25]

However, I have insisted throughout this paper that Frege's proposal is of meta-semantic nature (although it is in only in this sub-section that I have explicitly labeled it as such). I am therefore under no pressure to say that (T-Red) gives the meaning of an identity statement.

[24] See [6], §64.
[25] See [7] and [20].

I can instead maintain that the only semantically generated truth-condition for an identity statement is the trivial one (T-Triv). What Frege proposes is rather an account of what a subject's understanding of an identity statement consists in. And as we have seen, this account involves the principle for the identity of their semantic values:

(SV) $$[\![\mathbf{a}]\!] = [\![\mathbf{b}]\!] \leftrightarrow u \approx v$$

When this principle (SV) is combined with the trivial truth-condition (T-Triv), we do indeed get the reductive one (T-Red), which now emerges as a hybrid of semantic and meta-semantic facts.

A worry related to the one about compositionality concerns the compatibility of my Fregean account of reference with the semantic thesis that names and their mental counterparts are rigid designators. Consider a representation **a** associated with a referential attempt $\langle u, \approx \rangle$. Let f be the function determined from $\approx$ in accordance with the Principle of Individuation (PI). I have argued that **a** refers, if at all, to the object $f(u)$. One may then wonder whether my view isn't committed to some version of the now widely rejected *descriptivist* view of names, which identifies the semantic value of a name with that of some description associated with the name. Specifically, am I not committed to identifying the meaning of **a** with that of the description "the f of u," with the result that **a** isn't a rigid designator after all? (For instance, a chunk of physical matter may be part of different bodies in different possible worlds.)

The above discussion provides the resources needed to dismiss this worry. On my proposal, the nature of the function-argument structure $f(u)$ is entirely meta-semantic. The semantic value of **a**, if any, will be *the object* $f(u)$. How this referent is determined is entirely a meta-semantic matter, and thus of no immediate semantic significance. As far as semantics is concerned, **a** is a simple term or representation whose semantic value is just an object.

In fact, this view of the nature of the function-argument structure $f(u)$ enjoys independent evidence. Semantic structure is by and large accessible to consciousness; otherwise we wouldn't know or be rationally responsible for what we say and think. But someone can understand reference to shapes and bodies without having any conscious knowledge of how such reference is structured. Someone's competence with this structure may be located wholly at a subpersonal level. This is evidence that the structure isn't semantic. And if that is right—that is, if the function-argument structure of my account of reference isn't of semantic nature—then my account will be fully compatible with the rigidity thesis and in no danger of collapsing back into descriptivism (which is a semantic thesis).

2.4 Existence and uniqueness of referents

I now turn to two final worries about principles of individuation. Let I be a class of presentations and $\approx$ a unity relation defined on I. Assume we stand in an adequate relation to both the presentations and the unity relation. Then I am committed to the claim that I and $\approx$ implicitly define a partial function f such that

(PI) $$f(u) = f(v) \leftrightarrow u \approx v$$

and such that f is defined on all presentations in the field of the relation $\approx$.

Two worries arise concerning this view. Firstly, how do we know that there exist objects of the sort that are supposed to be in the range of the partial function f? For instance, assuming that lines exist and that each is parallel to itself, how do we know that their directions exist? To answer such skeptical questions, we need to analyze claims to the effect that Fs exist, where F is the concept implicitly defined by the class of presentations I and the unity relation $\approx$. These claims come out true provided there are semantic values falling under the concept F. And a semantic value falls under the concept F just in case it can be presented by a referential attempt $\langle u, \approx \rangle$ which is successful in the sense that $u \approx u$. But this is just what we have assumed. So the desired kind of semantic values *will* exist, and the claim that Fs exist will therefore be true.

It may be objected that this response trivializes ontological questions. How can what I have just described be all that is required for an object to exist? I admit that some of the objects I have talked about are very "light-weight," in the sense that their existence doesn't amount to very much. (I will have more to say about this in the next section, where I discuss mathematical objects.) But other objects are less "light-weight." For instance, the condition for a referential attempt to pick out a physical body is what we would expect: the presentations u must be a chunk of solid stuff, and this chunk must be related to other chunks in such a way as to define a unified whole of solid stuff that moves as a unit.

It may also be objected that my response to the first worry opens the door to all kinds of strange and unusual objects which we normally never talk about. For instance, let the class of presentations consist of people and the unity relation be siblinghood. On my account this is a coherent form of reference. Does this show that my account is committed to an extravagant and implausible ontology? Although I admit that my account is committed to a generous ontology, I believe it also has the resources needed to explain why this ontology is harmless. In particular, the apparent implausibility of the unusual objects which my account countenances can be adequately explained by the fact that canonical reference to such objects is very different from any kind of reference in which we ordinarily engage.

The second worry is that there may be *more than one* equally good candidates for playing the role of the function f. After all, whenever a function f satisfies (PI), then so does $\sigma \circ f$, where σ is any permutation of the Fs. Can we point to some feature of the intended function f that distinguishes it from its unintended rivals? Unfortunately, a more explicit characterization of the function f is out of the question. For a more explicit characterization—whether in the form of an algorithm or simply as a set of ordered pairs—would presuppose that the objects in the range of this function can be referred to in some way *other* than by means of this function. But on my account, there is no such alternative way of referring to these objects, since the function f plays an essential role in canonical reference to objects of the sort in question. So the function f cannot be given a more explicit characterization than that given by (PI).

Fortunately, this very fact can also be used to assuage the worry. Assume someone successfully uses a referential attempt $\langle u, \approx \rangle$. Assume some philosopher studies this referential attempt and asks which of two candidate referents this referential attempt has singled out. How can this question be answered? If my theory is completely general, the philosopher must herself pick out the two candidate referents by means of referential attempts. Since two referential attempts can pick out the same referent only if their unity relations are compatible, we may assume these referential attempts to be of the form $\langle v_i, \approx \rangle$. But then our philosopher can determine which candidate is right by determining whether $u \approx v_i$ for either i.

3 Referring to the natural numbers

I will now apply some of the ideas of the previous sections in an attempt to explain how we manage to refer to the natural numbers (where by 'we' I will mean ordinary speakers of English with at least basic competence in arithmetic). Since this is an account of an ability that we have, it will be relevant to consider some psychological facts about us. My account is based on two claims about our ordinary thought and talk about natural numbers. I begin by explaining and defending these two claims.

My first claim is that we regard the natural numbers as finite *ordinals*, individuated by their position in a well-ordering, rather than as finite *cardinals*, individuated by the cardinality of the sets or concepts whose numbers they are.[26] I grant that *one* way of thinking and talking

[26]The view that the natural numbers are finite cardinals is defended by the classical logicists Frege and Russell, as well as by contemporary neo-logicists. Although this view is compatible with my Fregean account of reference, I don't think it gives an accurate description of our ordinary arithmetical thought and talk.

about the natural numbers is by means of expressions of the form 'the number of Fs', as the cardinal-based approach will have it. What I deny is that this is our most fundamental or direct way of thinking and talking about the natural numbers. I offer four arguments for this claim.

1 *A phenomenological argument.* If there are seven apples on a table, we can think of the number seven as the number of apples on the table. Or (following Frege) we can think of seven as the number of numbers less than or equal to six. But neither or these ways of thinking of the number seven *feels* very direct or explicit. A more direct and explicit way of thinking of seven is by means of the standard numeral '7' or by means of any other numeral occupying the seventh position in a system of numerals with which we are familiar.

2 *An argument from the philosophy of language.* Expressions of the form 'the number of Fs' aren't singular terms but definite descriptions. It is for instance easily seen that such expressions aren't rigid (although the ones couched entirely in mathematical language will be *de facto* rigid, in the sense that the description is true of the same object in all possible worlds). But it is doubtful that there can be a kind of objects to which reference is possible *only* by means of definite descriptions. In order to refer to an object by means of a definite description, there must be some other, more direct way of thinking of the object.

3 *An argument from the number zero.* If our most fundamental way of thinking about the natural numbers is as cardinals, then zero would have been the most obvious and immediate number of them all. If on the other hand our most fundamental way of thinking about the natural numbers is as ordinals, then zero would be no more obvious or immediate than the negative numbers. This latter hypothesis accords much better with the late stage at which zero was admitted into mathematics as a number in good standing.

4 *A technical argument.* In [11] I showed that on the cardinal-theoretic approach, the standard proofs of some basic predicative arithmetical truths require impredicative second-order comprehension. This is not only unnatural but makes these very elementary arithmetical truths depend on much too theoretical principles. This problem goes away when the natural numbers are regarded as ordinals.

My second claim is that it is part of ordinary arithmetical competence that the natural numbers are *notation independent*, in the sense that they can be denoted by different systems of numerals. Indeed, even people with very rudimentary knowledge of arithmetic know that the natural numbers can be denoted by ordinary decimal numerals, by their counterparts in written and spoken English (and in other natural languages), and by sequences of strokes (perhaps grouped in fives). Many people also know alternative systems of numerals such as the Roman numerals and the numerals of position systems with bases other than ten, such as binary and hexadecimal numerals. To accommodate such alternative systems of numerals, we need a *general* condition for two numerals to denote the same number and for one numeral to be related to another in such a way that the number denoted by the former immediately precedes the number denoted by the latter.

I will here take a numeral to be any object that occupies a position in a well-ordering. In fact, since it is convenient to make the well-ordering explicit, I will take a numeral to be an ordered pair $\langle u, R\rangle$, where u is the numeral proper and R is the well-ordering in which u occupies a position. On this very liberal view of the matter, the numeral proper u need not be a syntactical object, at least not in any traditional sense. For instance, if a pre-historic shepherd counts his sheep by matching them with cuts in a stick, then these cuts count as numerals. Moreover, since R can be any well-ordering, these numerals refer to ordinal numbers but not necessarily to finite ones.

Our first task is to describe the equivalence relation that holds between two such pairs $\langle u, R\rangle$ and $\langle u', R'\rangle$ when they determine the same number. This equivalence relation must obviously be a matter of the two objects u and u' occupying analogous positions in their respective orderings, for instance, that both occupy the 17th position. More formally, $\langle u, R\rangle$ and $\langle u', R'\rangle$ are equivalent just in case there exists a relation C which

- is an order-preserving correlation of initial segments of R and R'
- is extensive enough to have both u and u' in its field
- is such that $C(u, u')$.[27]

Let $\langle u, R\rangle \approx \langle u', R'\rangle$ symbolize that the two ordered pairs are equivalent in this sense. Ordinal numbers are then individuated by the following abstraction principle:

(Id-N) $$N\langle u, R\rangle = N\langle u', R'\rangle \leftrightarrow \langle u, R\rangle \approx \langle u', R'\rangle$$

Note that the numbers to which the numerals are mapped are not equivalence classes of numerals but form their own category of objects. (Likewise, physical bodies are *sui generis* objects rather than equivalence classes of chunks of physical stuff.) Let '$O(x)$' be a predicate that holds of all and only the objects that can be presented in this way.

[27]The well-foundedness of the relations R and R' guarantees that there is an order-preserving correlation of initial segments extensive enough to include at least one of u and u'. And this is all we need to determine whether the condition in the main text is met.

Next we define a predecessor relation $P^{\#}$ on such pairs by letting $P^{\#}(\langle u, R\rangle, \langle u', R'\rangle)$ just in case there is an R'-predecessor v of u' such that $\langle u, R\rangle \approx \langle v, R'\rangle$. It is easily verified that $P^{\#}$ is a congruence with respect to $\approx$ (that is, that $P^{\#}$ doesn't distinguish between $\approx$-equivalent objects). This means that $P^{\#}$ induces a predecessor relation P on the ordinal numbers themselves, defined by

(Def-P) $$P(N\langle u, R\rangle, N\langle u', R'\rangle) \leftrightarrow P^{\#}(\langle u, R\rangle, \langle u', R'\rangle).$$

Finally, following the practice of ordinary counting, we let 1 be the first number. For instance, 1 may be presented as $N\langle \text{‘1'}, D\rangle$, where D is the familiar well-ordering of decimal numerals.

These definitions enable us to establish some of the basic axioms for ordinal numbers. We begin with the easier ones.

(O1) $O(1)$

(O2) $\neg\exists x\, P(x, 1)$

(O3) $P(x, y) \wedge P(x', y) \rightarrow x = x'$

(O4) $P(x, y) \wedge P(x, y') \rightarrow y = y'$

(O5) $\forall X[\exists x(Ox \wedge Xx) \rightarrow \exists x(Ox \wedge Xx \wedge \forall y((Oy \wedge Xy \wedge x \neq y) \rightarrow P(x, y)))]$

The proof of (O1) is trivial. For (O2), let n be any ordinal number. Then there is some presentation $\langle u, R\rangle$ such that $n = N(\langle u, R\rangle)$. If $P(n, 1)$, then u would come before the first element in a well-ordering. But this is impossible. (O3) follows from the observation that any two numerals which immediately precede a third are equivalent. (O4) follows from the observation that any two numerals which immediately succeed a third are equivalent. (O5) follows from the observation that the numerals are well-ordered.

However, I have not yet said anything very substantial about how many ordinals there are. For the purpose of describing the natural numbers (which I have identified with the finite ordinals), the only such principle we need is *the Successor Axiom*:

(O6) $\forall x(O(x) \rightarrow \exists y P(x, y))$

The Successor Axiom would follow immediately from a corresponding principle about presentations:

$$\forall\langle u, R\rangle\, \exists\langle u', R'\rangle\ P^{\#}(\langle u, R\rangle, \langle u', R'\rangle)$$

However, it is doubtful that this principle about presentations has the required epistemological status. I therefore instead adopt the following weaker principle:

$$\Box\forall\langle u, R\rangle\Diamond\exists\langle u', R'\rangle \; P^{\#}(\langle u, R\rangle, \langle u', R'\rangle)$$

This modified principle is extremely plausible. For assume we're given a presentation $\langle u, R\rangle$. Then it is possible that there should be some object u' not among the relata of R. Let R' be the result of adding the pair $\langle u, u'\rangle$ to the initial segment of R ending with u. Then $\langle u', R'\rangle$ is as desired. Moreover, combined with a claim to the effect that ordinals exist by necessity, this modified implies the Successor Axiom.

Finally, we need to specify some condition of finitude with which to restrict the ordinals such that we get all and only the natural numbers. I claim that this condition is simply that mathematical induction be valid of the natural numbers. That is, an ordinal n is a natural number (in symbols: $\mathbb{N}n$) just in case the following open-ended schema holds:

(MI) $\phi(1) \wedge \forall x \forall y[\phi(x) \wedge P(x, y) \rightarrow \phi(y)] \rightarrow \phi(n)$

Our characterization of the natural numbers has thus allowed us to prove all the Dedekind-Peano axioms.

4 The metaphysical status of the natural numbers

Does the account of reference to the natural numbers that I have outlined tell us anything about their metaphysical status? In particular, does the account favor either a platonist or a nominalist interpretation of the language of arithmetic?

Before addressing the question of platonism directly, it is useful to explain a fundamental difference between physical bodies and natural numbers having to do with the ways in which they possess properties. Consider the question whether a physical body x has some property, say being round. To answer this question, it's not sufficient to consider any proper part of x. Whether a body is round isn't determined by any of its proper parts but information is needed about the entire body. And there is nothing unusual about this case. It is in general true that, in order to determine whether a body x has some property G, one needs information about *many* parts of x. The question whether a body has some property G cannot in general be reduced to a question about any *one* of its proper parts. This means that a body can have properties in an irreducible way, that is, in a way that isn't reflected in any properties of any one of its proper parts

The situation is very different with natural numbers. Consider the question whether a natural number n has some mathematical property G, say the property of being even. In this case a standard presentation of n by some numeral (say a standard decimal numeral $\mathbf{n}$) suffices to answer the question. The question whether the natural number n possesses the property G can be reduced to a question about the numeral $\mathbf{n}$ by which n is presented. Here's why. All the usual arithmetical properties are definable (in second-order logic) from the predecessor relation P. And as (Def-P) shows, the question whether P holds between two natural numbers is itself reducible to the question whether the relation $P^{\#}$ holds between numerals. Natural numbers are therefore "impoverished" compared to numerals. For whenever a natural number n possesses some property, its doing so is inherited from the fact that the numerals that present n possess some associated property. Natural numbers are therefore "thinner" than the numerals that present them. This opens for a form of reductionism about natural numbers: Questions about such objects can be reduced to questions about their presentations.[28]

Given this reductionism, does it still make sense to say that numerals refer to natural numbers? I believe this question is best understood as the question whether it still make sense to ascribe semantic values to numerals. I will now argue that this does still make sense. One observation that supports this claim is the following. It is always the default assumption that a syntactically uniform class of expressions, such as the class of singular terms, should have a uniform semantic analysis. Now, when we analyze English and the language of arithmetic, singular terms such as '5' and '1001' seem to function just like terms such as 'Alice' and 'Bob'. The default assumption is therefore that all these terms function in similar ways. Since singular terms such as 'Alice' and 'Bob' clearly have semantic values (namely the physical bodies that they refer to), this provides at least some reason to think that arithmetical singular terms such as '5' and '1001' have semantic values as well.

It may be objected that this default assumption is overridden by our discovery that questions about natural numbers can be reduced to questions about the associated numerals. Since this reduction shows that it suffices to talk about the numerals themselves, the objection continues, there is no need to ascribe any sort of semantic values to numerals. However, this objection will succeed *only if the structure responsible for the reduction that we have discovered is also the kind of structure that matters for semantic analysis.* But I argued in Section 2.3 that this condition is not met, because the structure responsible for the reduction isn't semantically accessible but is entirely of meta-semantic nature.

If these arguments are on the right track, then it *does* make sense to ascribe semantic

[28]For reasons that I cannot here go into, I think that the same hold of other mathematical objects as well.

values to numerals. It is therefore perfectly true to say that natural numbers exist. So clearly my analysis isn't a nominalist one. However, my analysis is also far removed from a traditional platonist conception of mathematics. For although I have argued that numerals do have abstract semantic values, I have also argued that this fact allows of a deeper, non-semantic analysis on which any such reference to mathematical objects disappears. To ascribe to a language a certain semantic structure is to ascribe to it a certain kind of pattern of regularities in the ways truth-values of sentences are determined—a pattern of a kind that exists throughout human language and thought, and which has some important linguistic and psychological features. But this pattern may in turn rest on a deeper pattern on which reference to mathematical objects drops out. Reductionism about reference can therefore be incorrect on a semantic analysis, although correct on a different, more thoroughgoing form of analysis.

Where does this leave the question of mathematical platonism? If by 'mathematical platonism' is meant simply the view that there are true sentences some of whose semantic values are abstract, then my view is obviously a platonist one. But given how light-weight these semantic values are, this may be more of a reason to reject the above definition of mathematical platonism than for hard-line platonists to declare victory.

References

[1] George Boolos. To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables). *Journal of Philosophy*, 81(8):430–449, 1984. Reprinted in [**?**].

[2] Michael Dummett. *Frege: Philosophy of Language*. Harvard University Press, Cambridge, MA, second edition, 1981.

[3] Michael Dummett. *Frege: Philosophy of Mathematics*. Harvard University Press, Cambridge, MA, 1991.

[4] Gareth Evans. *Varieties of Reference*. Oxford University Press, Oxford, 1982.

[5] Kit Fine. *The Limits of Abstraction*. Oxford University Press, Oxford, 2002.

[6] Gottlob Frege. *Foundations of Arithmetic*. Blackwell, Oxford, 1953. Transl. by J.L. Austin.

[7] Bob Hale. *Grundlagen* §64. *Proceedings of the Aristotelian Society*, 97(3):243–61, 1997. Reprinted with a postscript in [8].

[8] Bob Hale and Crispin Wright. *Reason's Proper Study.* Clarendon, Oxford, 2001.

[9] Øystein Linnebo. Frege's Proof of Referentiality. *Notre Dame Journal of Formal Logic*, 45(2):73–98, 2004.

[10] Øystein Linnebo. Plural Quantification, 2004. In *Stanford Encyclopedia of Philosophy*, available at http://plato.stanford.edu/entries/plural-quant/.

[11] Øystein Linnebo. Predicative Fragments of Frege Arithmetic. *Bulletin of Symbolic Logic*, 10(2):153–74, 2004.

[12] Øystein Linnebo. To Be Is to Be an *F*. *Dialectica*, 59(2):201–222, 2005.

[13] Øystein Linnebo. Sets, Properties, and Unrestricted Quantification. In Agustín Rayo and Gabriel Uzquiano, editors, *Unrestricted Quantification: New Essays*, pages 149–178. Oxford University Press, Oxford, 2006.

[14] Charles Parsons. Ontology and mathematics. *Philosophical Review*, 80:151–176, 1971. Reprinted in [15].

[15] Charles Parsons. *Mathematics in Philosophy.* Cornell University Press, Ithaca, NY, 1983.

[16] John Perry. The Two Faces of Identity. In *Identity, Personal Identity, and the Self.* Hackett, Indianapolis, IN, 2002.

[17] Robert Stalnaker. On Considering a Possible World as Actual. *Proceedings of the Aristotelian Society*, Suppl. vol. 65:141–156, 2001.

[18] Kai Wehmeier. Consistent Fragments of *Grundgesetze* and the Existence of Non-Logical Objects. *Synthese*, 121:309–28, 1999.

[19] Timothy Williamson. *Identity and Discrimination.* Blackwell, Oxford, 1990.

[20] Crispin Wright. The Philosophical Significance of Frege's Theorem. In Richard Heck, editor, *Language, Thought, and Logic. Essays in Honour of Michael Dummett.* Clarendon, Oxford, 1997. Reprinted in [8].

[21] Crispin Wright. Response to Michael Dummett. In Matthias Schirn, editor, *Philosophy of Mathematics Today.* Clarendon, Oxford, 1998.

[22] Crispin Wright. The Harmless Impredicativity of $N^=$ (Hume's Principle). In Matthias Schirn, editor, *Philosophy of Mathematics Today.* Clarendon, Oxford, 1998.