

A Higher-Order Solution to the Problem of the Concept *Horse**

Nicholas K. Jones

This is a draft of a paper forthcoming in *Ergo*; please cite only the final version.

Abstract

This paper uses the resources of higher-order logic to articulate a Fregean conception of predicate reference, and of word-world relations more generally, that is immune to the concept *horse* problem. A prominent style of expressibility problem for views of broadly this kind, versions of which are due to Linnebo, Hale, and Wright, is then addressed.

Central to an account of the relationship between language and reality is an account of that between predicates and reality. Frege's problem of the concept *horse* (§1) threatens one prominent style of account, on which predicates refer to extra-linguistic entities fundamentally different in kind from those to which singular term refer. Here, I explore the interaction between this classic problem in the analytic tradition's history and a contemporary view about the interpretation of higher-order logic, developing a higher-order Fregean semantics and metaphysics that is immune to the problem. At the heart of my approach is a seemingly anti-Fregean semantic framework that appears to withhold reference from predicates. I argue (§2) that this appearance is misleading: from a higher-order perspective, this seemingly anti-Fregean semantics can be seen to treat predicates referentially. After articulating the conception of higher-order logic on which this argument depends (§3), I show how to combine this approach with other central principles of Fregean semantics and metaphysics without giving rise to the concept *horse* problem (§4). I then (§5) address a prominent style of objection to views of broadly this kind, versions of which are due to Øystein Linnebo, Bob Hale, and Crispin Wright. The lesson is that the concept *horse* problem for Fregean conceptions of predicate reference—and of the language-reality relationship more generally—and their apparent conflict with non-referential treatments of predicates, both stem from a failure to take seriously the existential import of higher-order quantification alongside unwarranted insistence on formulating Fregean semantics in first-order terms.

*Thanks to Bill Brewer, Craig French, Anil Gomes, David Liggins, Ian Phillips, Scott Sturgeon, Lee Walters, Caspar Wilson, and Al Wilson for comments and discussion. Thanks also to an audience at KCL for their feedback. An Area Editor and anonymous referee for *Ergo* offered some very helpful suggestions.

To be clear, my goal is not scholarly reconstruction of Frege’s view. Rather, my goal is to articulate and explore a recognisably Fregean position that captures the core of his conception of predicate reference. Although I do think my proposal captures much of what Frege intended, I won’t argue for that here. I leave it to the reader to decide the extent to which my proposal deserves the label ‘Fregean’.

1 The Problem of the Concept *Horse*

This section outlines a Fregean conception of predicate reference, and its associated problem of the concept *horse*.

1.1 The Problem

How does the semantic function of predicates connect them with extra-linguistic reality? Our Fregean answer embodies two central claims. Firstly, just as semantically non-defective singular terms refer to constituents of (typically) extra-linguistic reality, so too do semantically non-defective predicates.¹ Secondly, expressions from different semantic categories pick out entities from different ontological categories, so that (at least some) ontological distinctions can be read off from the semantic structure of language. Combining these two claims, we arrive at the following picture of the relationship between language and reality. Singular terms refer to entities from a certain ontological category: *object*. Predicates refer to entities from a fundamentally different ontological category: *concept* (aka *property*). Frege glossed the difference between these categories by saying that whereas objects are “complete”, “self-standing” and “saturated”, concepts are “incomplete” and “unsaturated”. These Fregean metaphors will play no role in my discussion.

We can regiment this *Fregean theory of predication* as the conjunction of:

Reference: Predicates refer.

Objecthood: Whatever singular terms may in principle refer to is an object; nothing else is an object.²

Concepthood: Whatever predicates may in principle refer to is a concept; nothing else is a concept.

Exclusivity: No object is a concept.

¹ I leave the restriction to semantically non-defective expressions tacit henceforth, and likewise for the “typically” qualification needed to accommodate talk about language. I also ignore the possibility of semantically non-defective yet non-referring expressions, putative examples of which include ‘Santa’, ‘Zeus’ and ‘Mary Poppins’.

² I use ‘may in principle’ to abstract away from the merely practical and cognitive linguistic limitations of finite beings like ourselves. Read ‘a singular term may in principle refer to *y*’ as roughly: singular terms are not precluded from referring to *y* by their semantic role alone. A singular term may in principle refer to *y* despite it being impossible for any such term to refer to *y*; for the impossibility may flow not from the semantic role of singular terms alone, but also from the impossibility of, say, cognitive contact between language users and *y*. (Hale, 2013, ch1, 193–195) employs a similar but more restrictive modality in his Fregean explication of objecthood.

It is not hard to derive a contradiction from these principles. The problem of the concept *horse* is the problem of resolving this apparent inconsistency in the Fregean theory of predication.

Consider the paradigmatically non-defective predicate ‘est un cheval’. By **Reference** and **Concepthood**: ‘est un cheval’ refers to a concept. Which concept? Our Fregean answer is line (1) of this *concept horse argument*:

- (1) Premiss: The predicate ‘est un cheval’ refers to the concept *horse*.
- (2) Premiss: ‘The concept *horse*’ is a singular term.
- (3) Premiss: ‘The concept *horse*’ refers to the concept *horse*.
- (4) By (1), **Concepthood**: The concept *horse* is a concept.
- (5) By (2), **Objecthood**: Whatever ‘the concept *horse*’ refers to is an object.
- (6) By (3), (5): The concept *horse* is an object.
- (7) By (6), **Exclusivity**: The concept *horse* is not a concept.
- (8) By (4), (7): The concept *horse* both is and is not a concept.

Since (8) is inconsistent, either the Fregean theory of predication or one of (1)–(3) is false. But which? The next subsection discusses the motivations for **Reference**, **Objecthood**, **Concepthood**, **Exclusivity**, (1), (2), and (3).

1.2 Motivating the Premisses

This subsection motivates the concept *horse* argument’s premisses. The point is to show that it isn’t merely a historical curiosity, but a live problem arising from independently attractive theses about how language functions.

Two views about **Objecthood** and **Concepthood** are available. The first sees them as definitional stipulations settling what Fregeans mean by ‘object’ and ‘concept’. Those words are intended to pick out certain classes of thing. On this approach, Fregeans use semantic vocabulary to determine which classes. Since that’s clearly legitimate, **Objecthood** and **Concepthood** should be uncontentious in the present context.

The other view about **Objecthood** and **Concepthood** begins from the idea that the ontological categories of *object* and *concept*—aka *property*—are reasonably well understood notions with important theoretical work to do. The question then arises: what is it to belong to those categories? **Objecthood** and **Concepthood** can be seen as substantive answers to that question, to be evaluated on their theoretical merits as two potential options amongst many. The goal is to use well understood logico-linguistic notions to elucidate some more elusive ontological ones, perhaps in service of what Bob Hale (2013, 17) calls a “Fregean approach to ontological questions”: “ontological categorization ... is dependent upon and derivative from prior logical categorization of expressions” (Hale, 2013, 10).

This should be attractive insofar as semantic notions are well understood, whereas ontological ones are elusive yet theoretically important.³ We needn't decide between this and the first, definitional, view of **Objecthood** and **Concepthood** here.

Exclusivity should seem more contentious. Given what our Fregean means by 'object' and 'concept'—at least, on the first (definitional) view of **Objecthood** and **Concepthood**—it amounts to the thesis that no singular term and predicate could possibly co-refer. That's hardly a neutral starting point for an inquiry into predicate reference, and the relationship between semantic and ontological categories. So what might be said in its favour?

Exclusivity follows from two theses.⁴ I'll say something in favour of each, though I cannot offer a full defence here. The point I want to make is that **Exclusivity** follows from a seemingly coherent conception of reference, and of the semantic differences between singular terms and predicates. The problem of the concept *horse* might be taken to undermine that combination of views. The Fregean framework I developed below shows that that conclusion is not inevitable.

The first thesis connects reference with semantic role:

Determination: The semantic properties of each expression are fully determined by its referent.^{5,6}

The idea is that an expression's semantic properties are inherited from its referent. Or alternatively, and more strongly: the point of assigning referents to expressions is, in part, to codify their semantic properties; that's part of reference's core theoretical role.⁷ In a slogan: semantics is exhausted by reference. It follows (i) that co-reference with an expression e suffices for functioning semantically just like e , and hence also (ii) that referring to an object (concept) suffices for functioning semantically as a singular term (predicate).⁸

³ Of course, **Objecthood** and **Concepthood** are not the only possible semantic analyses of ontological categories. See (Hale, 2013, 31–34) for another.

⁴ Frege (1892, 193) uses the unity of the proposition to motivate **Exclusivity**. Since that problem is highly obscure, I won't consider it here.

⁵ (MacBride, 2011) resolves the problem by, in effect, rejecting **Determination**.

⁶ **Determination** is a cousin of Crispin Wright's (1998, 73) Reference Principle (RP): if expressions e, e^* are co-referential, then substitution of e for e^* preserves grammaticality in all contexts, and preserves truth-value in all extensional contexts. Since the exact relationship between RP and **Determination** is somewhat complicated, I won't examine it properly here. Moreover, the theoretical basis for RP is unclear. It entails that expressions with the same referent belong to the same grammatical category. Why should reference constrain syntax in that way? I see no obvious reason why it should; and (Schiffer, 2003, 93) and (Oliver, 2005) argue that it doesn't, though see (Trueman, MSb) for a defence of RP based on a more sophisticated notion of substitution. My own view is that **Determination** captures whatever theoretical support RP may enjoy, by replacing RP's route from reference to syntax with one from reference to semantics. Note also that **Determination** is neutral about how rich the referents of expressions must be. For example, an extensionally individuated conception of predicate referents is inadequate for modal languages, where grammaticality-preserving substitution of coextensive predicates can affect truth-value. That shows that Frege's extensional conception of predicate referents was inadequate for modal languages. It does not show that **Determination** is flawed, since an intensionally individuated conception of predicate referents is adequate for modal languages.

⁷ Dummett (1973, 181, 190–191, 210 and elsewhere) discusses the conception of reference as semantic role.

⁸ Suppose expression e refers to object o . By **Objecthood**: e is co-referential with some potential term t .

The second thesis is about the semantic capacities of singular terms and predicates:

Linguistic Exclusivity: No expression can function semantically as both a singular term and predicate.⁹

I do not know how to argue for this thesis on general theoretical grounds. It is, however, supported by the apparent absence of counterexamples from natural language, and our apparent inability to invent any.¹⁰ The best explanation for this absence of counterexamples is their impossibility, arising from incompatibility between the semantic functions of singular terms and predicates.¹¹

Determination and **Linguistic Exclusivity** entail (**Ontological**) **Exclusivity**. Suppose for *reductio* that *o* is both an object and a concept. Because *o* is an object, **Objecthood** implies that a singular term could in principle refer to *o*. Because *o* is a concept, **Determination** entails that any such term would also function semantically as a predicate, contrary to **Linguistic Exclusivity**. So by *reductio* and since *o* was arbitrary: nothing is both an object and a concept; **Exclusivity** holds. The conception of reference that underwrites **Determination** thereby converts the exclusivity of the linguistic distinction between singular term and predicate into an exclusive ontological distinction between objects and concepts, the referents of singular terms and predicates.

Think of premiss (1) as a placeholder for whatever specification the Fregean ultimately gives of the referent of ‘est un cheval’. Rejecting this particular proposal does not block the argument. One must reject every principle that uses a singular term to specify the referent of ‘est un cheval’.¹² One must therefore also reject:

(1 \exists) $\exists x$ (‘est un cheval’ refers to *x*).

Were (1 \exists) true, one could introduce a new term *t* for whatever witnesses its initial existential quantifier. The argument could then be run just as before, substituting *t* for ‘the concept *horse*’ throughout. So rejection of (1) blocks the argument only in conjunction with rejection of (1 \exists). Yet that appears to contradict **Reference**; for if predicates refer, then ‘est un cheval’ refers, and so it refers to something, which is just what (1 \exists) says. By

By **Determination**: *e* functions semantically just like *t*. So if *e* refers to an object, *e* functions semantically as a term. Likewise *mutatis mutandis* for concepts and predicates.

⁹ Qualification: only ambiguous expressions can function semantically as both singular term and predicate, with those different semantic functions corresponding to different disambiguations. Alternatively: no semantically individuated expression can function semantically as both singular term and predicate.

¹⁰ Frege considers and rebuts the best apparent examples in (Frege, 1892, 183–184).

¹¹ I say that **Linguistic Exclusivity** is the best explanation, not that it is the only explanation. One might instead appeal to incompatibility between the syntactic functions of singular terms and predicates. But surely syntax is up to us in a way that semantics is not. If I choose to use *e* so that it can grammatically occur in the positions of both singular terms and predicates, then surely it thereby does come to have the syntactic function of both singular terms and predicates in my idiolect. Whether both kinds of sentence are meaningful and express truth-conditions (under a single assignment of semantic properties to *e*) is a different matter entirely. Whereas we can choose what strings count as grammatical in our own idiolects, we cannot choose what truth-conditions there are to express, or how they can be compositionally generated from the semantic properties of other expressions.

¹² See (Wright, 1998) and (MacBride, 2008, §4.2) for critical discussion of key attempts to specify the referents of predicates without using singular terms. My proposal is another attempt to do so.

way of a spoiler, my proposal provides a way to consistently combine rejection of (1 \exists) with acceptance of **Reference**.

Perhaps the most dubious premiss is (2); for ‘the concept horse’ is arguably a definite description, and it’s doubtful whether descriptions are singular terms. Three comments follow.

Firstly, ‘the concept *horse*’ is unlike paradigmatic definite descriptions like ‘the King of France’. Its referent isn’t determined as whichever individual satisfies the classificatory material following ‘the’. It’s more like ‘the mountain Everest’, which appears to be a singular term for Everest, which the term also indicates (in some sense) to be a mountain. If so, then ‘the concept *horse*’ is a singular term for *horse*, which it also indicates (in some sense) to be a concept.

Secondly, (1) instantiates (1 \exists) only if ‘the concept *horse*’ is a singular term. Since (1) was offered as a placeholder account of what ‘est un cheval’ refers to, it should instantiate (1 \exists). So ‘the concept *horse*’ should be a singular term. If it isn’t, introduce a new term *t* for whatever witnesses (1 \exists) and run the argument just as before, substituting *t* for ‘the concept *horse*’ throughout.

Thirdly, whether treating ‘the concept *horse*’ as a definite description blocks the concept *horse* argument depends on the semantics of definite descriptions. The most prominent accounts that differentiate descriptions from singular terms assimilate them to quantifiers binding variables in singular term position.¹³ These views leave the argument unaffected. If ‘the concept *horse*’ is a description of this kind and (1) is true, then we can introduce a new term *t* for whatever witnesses (1)’s description-quantifier.¹⁴ The argument can then be run just as before, substituting *t* for ‘the concept *horse*’ throughout. Absent an alternative semantic analysis of ‘the concept *horse*’, rejection of (2) won’t save the Fregean theory of predication. So I’ll continue to treat ‘the concept *horse*’ as a singular term.

Premiss (3) is no less plausible than any other instance of:

- *n* refers to α .

(Instances are obtained by replacing ‘*n*’ with a name for a singular term, and ‘ α ’ with a translation of the named term into our metalanguage.)

A translation of an expression is co-referential with it. So in any instance of this schema, the referent of the expression that replaces ‘*n*’ is co-referential with the expression that replaces ‘ α ’. So each instance of the schema is true. Since ‘the concept *horse*’, as used by me right now, translates ‘the concept *horse*’, as used by me right now, (3) is an instance of this schema. So (3) is true. Moreover, it is utterly mysterious what ‘the concept *horse*’ might refer to, as used by me right now, if not the concept *horse*.

We’ve seen that attractive and coherent semantic stories can be told in support of **Objecthood**, **Concepthood**, **Exclusivity**, (2), and (3). I have no similar story to offer in support of **Reference**, which is what motivates (1) as a placeholder account of what ‘est un

¹³ Wright (1998, 79) suggests a description-quantifier binding variables in predicate position. But then ‘the concept *horse*’ cannot grammatically occupy the second argument position of ‘refers to’, which requires a singular term, and (1) doesn’t instantiate (1 \exists).

¹⁴ Assumption: description-quantifiers have existential import, so \ulcorner (the $x : A$) $B\urcorner$ entails \ulcorner ($\exists x : A$) $B\urcorner$.

cheval' refers to.¹⁵ The best motivation I see for **Reference** is its role in an attractive account of how language functions and connects with extra-linguistic reality. **Reference** should therefore be attractive insofar as the overall Fregean logico-metaphysical package is. It is, however, tempting to regard the concept *horse* argument as refuting that package, and thereby undermining the motivation for **Reference**. My goal is to show how temptation may be resisted, by formulating a version of the Fregean theory of predication that does not motivate (1), and is consistent with the falsity of all reference clauses of (1)'s form. The next section lays the groundwork by arguing that an apparently non-referential semantics for predicates is really a referential semantics. Yet there is no role in that theory for principles like (1).

2 From Non-Referential Semantics to Referential Semantics

This section argues that an apparently non-referential semantic analysis of predicates is really a referential semantics compatible with rejection of (1) (and all other principles of its form).

A semantic analysis of a language's singular terms and predicates is adequate only if it combines with the language's compositional rules to assign truth-conditions to all atomic predications in the language. Focussing on the monadic case for simplicity, the natural compositional rule for atomic predication is:¹⁶

Atom: An atomic predication $\lceil \Phi(t) \rceil$ is true iff monadic predicate Φ applies to the referent of singular term t .

Given **Atom**, the semantic clauses for terms and predicates must jointly encode two kinds of information. Firstly, they must encode what the terms refer to. The simplest way of doing so is for the semantic clauses for terms to be principles like:

- (a) 'a' refers to Al.
- (b) 'b' refers to Beth.

I'll assume henceforth that the semantic clauses for singular terms take this form. Note that two singular terms flank 'refers to'.

Secondly, the clauses must encode what the predicates apply to. On the referential Fregean semantics driving the concept *horse* problem, the semantic clauses for predicates are principles like:

- (Fi) 'F' refers to the concept *eats cookies*.

¹⁵ What about the conception of reference as encoding semantic role that I used to motivate **Exclusivity**, via **Determination**? The difficulty is that there may be multiple ways to fix semantic properties, other than via assignments of extra-linguistic referents. The conception ensures that if an expression refers, its referent determines its semantic properties; but the conception alone doesn't ensure that all expressions refer.

¹⁶ World-relativised analogues of **Atom** and the other semantic principles that follow are needed for languages containing modal vocabulary. I ignore this complication in the main text.

(Gi) ‘*G*’ refers to the concept *generous*.

As with (a) and (b), two singular terms flank ‘refers to’. Premiss (1) of the concept *horse* argument is of this form. However, such clauses alone do not settle what ‘*F*’ and ‘*G*’ apply to. Further machinery is needed to connect predicate reference with applicability. The natural candidate is:

- For any monadic predicate Φ and object o , Φ applies to o iff o falls under the concept referred to by Φ .

The concept *horse* argument shows that this semantic analysis of predicates is inadequate, given **Objecthood**, **Concepthood**, **Exclusivity**, (2), and (3); for those principles—or suitable analogues in the case of (2) and (3)—are inconsistent with (1), (Fi), and (Gi).

A semantic theory needn’t follow this two-part specification of a referent for each predicate, plus an independent theory of application. The semantic clauses for predicates may instead specify their application conditions directly, as in:

(Fii) For any object o , ‘*F*’ applies to o iff o eats cookies.

(Gii) For any object o , ‘*G*’ applies to o iff o is generous.

Combined with (a) and (b), (Fii) and (Gii) imply e.g.:

- ‘*F*’ applies to the referent of ‘*a*’ iff Al eats cookies.
- ‘*G*’ applies to the referent of ‘*b*’ iff Beth is generous.

So compositional rule **Atom** assigns truth-conditions to ‘*F(a)*’ and ‘*G(b)*’ thus:

- ‘*F(a)*’ is true iff Al eats cookies.
- ‘*G(b)*’ is true iff Beth is generous.

Given matching clauses for all terms and predicates of the language, the result is a systematic assignment of truth-conditions to all its atomic predications.

Semantic theories of this kind appear to treat predicates non-referentially:¹⁷ (Fii) and (Gii) do not specify referents for ‘*F*’ and ‘*G*’, and there is no need for a separate theory of the relationship between predicate reference and applicability. Defenders of this approach can thus consistently reject (1), (Fi), (Gi), and all similar principles that use singular terms to specify the referents of predicates. Such principles play no role in semantic theories of this kind, rendering them immune to the concept *horse* problem. Yet this comes at a cost to the Fregean: it appears to conflict with **Reference**, and hence with the Fregean theory of predication. I’ll now argue that this appearance is misleading. Semantic theories employing principles like (Fii) and (Gii) really treat predicates referentially, even when coupled with rejection of (1), (Fi) and (Gi).

Why think of singular terms as referential under a semantics that employs clauses (a) and (b)? One might answer: because those clauses employ the notion of reference; they

¹⁷ Wright (1998, 76) regards clauses like (Fii) and (Gii) as non-referential, or at best only implicitly referential, and rejects them on those grounds. I’ll argue that Wright’s wrong to do so.

say what ‘*a*’ and ‘*b*’ refer to. Yet what does employing the notion of reference amount to? Although the word ‘refers’ appears in both (a) and (b), that’s neither here nor there. That word could have been used to mean all sorts of different things. And we could, if we wanted, reformulate **Atom**, (Fii) and (Gii) so that ‘refers’ appears in them too:

Atom*: An atomic predication $\lceil \Phi(t) \rceil$ is true iff monadic predicate Φ refers to the referent of singular term *t*.

(Fii*) For any object *o*, ‘*F*’ refers to *o* iff *o* eats cookies.

(Gii*) For any object *o*, ‘*G*’ refers to *o* iff *o* is generous.

That is a cheap way of making predicates referential. It renders debate about predicate reference merely terminological. A non-terminological debate requires an account of reference that isn’t tied to any particular mode of linguistic expression.

We obtain a better notion of referentiality by focussing on word-world relations. The idea is that an expression is referential when its semantic role involves a relation with some particular aspect of extra-linguistic reality, different aspects for expressions with different semantic roles. What it is for the expression to function semantically as it does is for it to bear such-and-such relation to such-and-such aspect of reality. The hallmark of such a relationship is the possibility of existential generalisation into the expression’s semantic clause (in an adequate semantic theory for the language). That semantic clause captures the expression’s semantic role; it says how the expression semantically behaves, which governs its interaction with other expressions, via the compositional rules, to determine the semantic properties of complex expressions in which it occurs. When existential generalisation into a semantic clause is possible, the relevant semantic role involves a relation with something, namely whatever was generalised out of the original clause. Whatever else that something may be, it’s certainly an aspect, or constituent, of reality. Hence:

Criterion of Referentiality: An expression is referential iff non-trivial existential generalisation into its semantic clause is possible.

The non-triviality restriction rules out two cases. (i) The term for *e* can always be existentially generalised out of *e*’s semantic clause, e.g.: ‘‘*a*’ refers to Al’ entails ‘ $\exists x(x$ refers to Al)’. (ii) One can existentially generalise a sentence without binding any variable within it, e.g.: ‘‘*a*’ refers to Al’ entails ‘ $\exists x($ ‘*a*’ refers to Al)’. Neither case suffices for an interesting word-world relation between ‘*a*’ and anything else.

By way of illustration, consider the case of singular terms. Semantic clauses like (a) and (b) are open to existential generalisation since they entail:

(a \exists) $\exists x($ ‘*a*’ refers to *x*).

(b \exists) $\exists x($ ‘*b*’ refers to *x*).

So by the **Criterion**: singular terms governed by such clauses are referential. The account correctly classifies singular terms as referential.

What about predicates? It depends on their semantic clauses. Clauses like (1), (Fi) and (Gi) are open to existential generalisation since they entail:

(I \exists) $\exists x$ (‘est un cheval’ refers to x).

(Fi \exists) $\exists x$ (‘ F refers to x ’).

(Gi \exists) $\exists x$ (‘ G refers to x ’).

So by the **Criterion**: predicates governed by such clauses are referential.

What about a semantics based instead around clauses like (Fii) and (Gii), or (Fii*) and (Gii*)? At a first pass, such a semantics appears non-referential. The only expressions in those clauses open to existential generalisation are ‘ F ’ and ‘ G ’ themselves. So by the **Criterion**: predicates governed by such clauses are not referential. Those clauses entail the existence of the predicates they concern, but not anything else to which they’re related. The **Criterion** thus appears to vindicate the assessment as non-referential of semantic analyses of predicates based around clauses like (Fii) and (Gii). As I’ll now argue, however, that appearance is mistaken.

First-order quantifiers bind variables occupying the positions of singular terms. The only expressions in (Fii) and (Gii) open to first-order existential generalisation are ‘ F ’ and ‘ G ’. *Second-order* quantifiers bind variables occupying the positions of predicates.¹⁸ In a second-order setting, ‘ F ’ and ‘ G ’ are not the only expressions in (Fii) and (Gii) open to existential generalisation. The predicates ‘eats cookies’ and ‘is generous’ are open to second-order existential generalisation, and (Fii) and (Gii) entail:¹⁹

(Fii \exists) $\exists X$ (for any object o , ‘ F ’ applies to o iff $X(o)$).

(Gii \exists) $\exists X$ (for any object o , ‘ G ’ applies to o iff $X(o)$).

So by the **Criterion of Referentiality**: predicates governed by semantic clauses like (Fii) and (Gii) are referential. Although their semantic role doesn’t involve a relation to any first-order aspects of reality, it does involve a relation to second-order aspects of reality.

We can rewrite some of these principles slightly, to clarify what’s going on. We replace quantifiers over objects with familiar first-order quantifiers. We also employ the following defined, mixed-order predicate:

- 2-REFERS(x, Y) =_{df.} $\forall z$ (x applies to z iff $Y(z)$).

Note that ‘2-REFERS’ takes a term or term-variable in its first argument position, and a predicate or predicate-variable in its second. In this respect it’s unlike the reference predicate in (1), (a), (b), (Fi), and (Gi), which takes a term or term-variable in each of its argument positions. To differentiate them, I’ll call the original notion 1-reference, and this newly defined notion 2-reference. My key theses are that: (i) 2-reference is the appropriate notion of reference for predicates; (ii) the **Criterion of Referentiality** classifies 2-reference

¹⁸ An excellent overview of second-order logic is (Shapiro, 1991).

¹⁹ Assumption: the following rule of second-order existential generalisation is valid: from A , one may infer $\lceil \exists V A^{V/\Phi} \rceil$, where $A^{V/\Phi}$ results from A by substituting zero or more occurrences of the monadic predicate Φ with the monadic second-order variable V , so that all and only those newly substituted in variables are bound by the initial quantifier in $\lceil \exists V A^{V/\Phi} \rceil$. This is standard amongst advocates of second-order logic, e.g. (Shapiro, 1991, 66), (Hale, 2013, 181).

as a genuine notion of reference because it expresses a relation between expressions and second-order reality, as revealed by the possibility of existential generalisation into the second argument position of '2-REFERS'; and (iii) a referential semantics for predicates formulated using 2-reference is immune to the concept *horse* problem because principles like (1) play no role in such a semantics, that is, because such a semantics employs 2-reference in place of 1-reference in its clauses for predicates.

Using this notation, the supposedly non-referential clauses (Fii) and (Gii) become:

(Fii) 2-REFERS('F', eats cookies).

(Gii) 2-REFERS('G', is generous).

Those entail:

(Fii \exists) $\exists X(2\text{-REFERS}('F', X))$.

(Gii \exists) $\exists X(2\text{-REFERS}('G', X))$.

The referential similarity between (a), (b), (Fi), and (Gi) on the one hand, and (Fii) and (Gii) on the other is now clear. Just as the former say that 'a', 'b', 'F' and 'G' 1-refer to certain first-order aspects of reality, the latter say that 'F' and 'G' 2-refer to certain second-order aspects of reality. The existence of such aspects is expressed by (a \exists), (b \exists), (Fii \exists), and (Gii \exists), which follow by existential generalisation from (a), (b), (Fii), and (Gii). The exact word-world relation differs between the two cases, as does the order of quantification needed to satisfy the **Criterion of Referentiality**. But some word-world relation is present in each case, and the **Criterion** classifies predicates governed by clauses like (Fii) and (Gii) as referential. The defined higher-order reference predicate '2-REFERS' allows us to combine this referential similarity with linguistic similarity, thereby making vivid (Fii)'s and (Gii)'s referential standing. Those semantic clauses specify the 2-referents of predicates explicitly; and they do so without using singular terms. The concept *horse* problem does not arise because the appropriate reference clause for 'est un cheval' is not (1) but:

- 2-REFERS('est un cheval', is a horse).

Note that since 'is a horse' is a predicate, this is well-formed.

I've argued that a seemingly non-referential semantics for predicates is revealed as a referential semantics, when viewed from a second-order perspective. The argument rests on the following assumption, which hasn't yet been made explicit: just as first-order existentially quantified sentences express existence claims, so too do second-order existentially quantified sentences; it's just that they express second-order existence, rather than first-order. Without that assumption, (Fii \exists) and (Gii \exists) would not capture the existence of (second-order) aspects of reality to which 'F' and 'G' 2-refer. The **Criterion of Referentiality** would then be misapplied in the case of second-order quantification, and predicates would not count as referential. What justifies this assumption? And what does second-order quantification mean? The next section addresses these questions. §4 then formulates a Fregean theory of predication in higher-order terms.²⁰

²⁰ After this paper had been accepted for publication, I encountered Rob Trueman's (MSa), the main

3 Second-Order Quantification and Existence

What does second-order quantification mean? What is the relationship between second-order quantification and existence? I cannot hope to settle these matters of major controversy here. In answer to the first question, I simply assume without argument a popular and attractive conception of second-order quantification and its relationship to the first-order. My primary concern is with second-order quantification's role in a Fregean theory of predication, given this conception of it, not a defence of that conception. In answer to the second question, I offer some preliminary justification for my preferred answer, though I won't pretend that my discussion is utterly decisive.

What do second-order quantifiers mean, and how do they relate to first-order quantifiers? My preferred answer is as follows.²¹

Second-order quantification is a perfectly legitimate and intelligible form of genuine (non-substitutional) quantification whose truth-conditions cannot be explicated in first-order terms. It is not a disguised form of first-order quantification, or quantification over a special kind of object in the range of unrestricted first-order quantification or to which singular terms may in principle refer. Second-order quantification is a *sui generis* form of (non-substitutional) quantification that must be understood in its own terms or not at all. There is thus a fundamental semantical distinction between these two orders of quantification: first- and second-order quantified sentences express fundamentally different kinds of quantificational truth-condition. The familiar classification of quantifiers into the substitutional and the objectual is therefore not exhaustive; for second-order quantification is neither substitutional nor objectual (since the objects constitute the range of unrestricted first-order quantification). The semantic role of second-order quantifiers is both genuinely (non-substitutionally) quantificational and yet irreducibly second-order.

This fundamental semantic distinction between the first- and second-order is essential to my proposal's success. To see why, suppose the truth-conditions of second-order quantified sentences can be explicated in first-order terms, using first-order quantification over some privileged range of objects. Then the second-order existential generalisations—i.e. (Fii \exists) and (Gii \exists)—of the reference-clauses for predicates—i.e. (Fii) and (Gii)—are semantically equivalent to first-order quantified sentences of ($1\exists$)'s form: $\exists xR$ ('est un cheval', x). This ' R '-predicate is then the appropriate reference predicate for predicates, the result of translating '2-REFERS' into the first-order language used to explicate the truth-conditions of second-order quantification. So a new singular term t could be introduced for the quantifier's witness, and the concept *horse* argument run just as before, substituting

ideas of which are in many ways very similar to those just outlined. Trueman and I both regard 2-reference as the appropriate notion of reference for predicates. Unlike me, however, Trueman has no role for the **Criterion of Referentiality**. He focusses instead on the fact that 1-reference and 2-reference both allow formulation of disquotational principles, using that to justify regarding both as genuine notions of reference. He also uses the impossibility of formulating disquotational principles for predicates in terms of 1-reference to argue against semantic clauses like (1), (Fi), and (Gi). I'd like to thank Trueman for some very interesting correspondence about these issues.

²¹ For discussion of this kind of approach, see: (Prior, 1971, ch3), (Rayo and Yablo, 2001), (Williamson, 2003, esp.§9), (Rayo and Williamson, 2003, §1), (Wright, 2007), (MacBride, 2008, §19.2.2), (Williamson, 2013, 254–261).

t for ‘the concept *horse*’ throughout. If second-order quantification is to save the Fregean theory of predication, a fundamental semantic distinction between the first- and second-order is required.

The best arguments for this semantic distinction employ variants of Russell’s Paradox and Cantor’s Theorem.²² The very argument Frege thought had undermined his philosophy of mathematics may thereby help to save his theory of predication from the concept *horse* problem, by motivating the conception of higher-order logic needed for a consistent Fregean theory of predication. The concept *horse* argument turns on blurring the semantic distinctions between expressions of different semantic types in a manner akin to Basic Law V.²³ Both Basic Law V and the problematic version of the Fregean theory of predication—i.e. the version employing semantic clauses like (1), (Fi), and (Gi)—require an object in the range of first-order quantification for each (possible) predicate, to serve as its referent/extension, different objects for non-coextensive predicates. Despite this common source, however, the nature of the problems is quite different. Whereas elementary logic reveals Basic Law V’s inconsistency, the concept *horse* problem also requires **Objecthood, Concepthood, Exclusivity**, (2), and (3).

Do second-order quantified sentences express existence claims? The argument from the **Criterion of Referentiality** to the referentiality of predicates governed by (Fii) and (Gii) presupposes that second-order formulae like:

$$(Fii\exists) \quad \exists X(2\text{-REFERS}('F', X))$$

$$(Gii\exists) \quad \exists X(2\text{-REFERS}('G', X))$$

do indeed express existence claims. So one might respond by denying that they do. Per-

²² See (Williamson, 2003, §9), (Rayo and Williamson, 2003, §1), (Linnebo, 2006). By way of illustration, here’s a quick version of one such argument. Assume that second-order quantification and unrestricted first-order quantification are both legitimate. Were second-order quantifiers in the same kind of semantic business as first-order quantifiers, a first-order account of what second-order quantifiers range over would be possible. But a version of Cantor’s Theorem shows that to be impossible. Such an account would require a first-order representative for every second-order “entity”. In an extensional setting, that minimally requires an f from the first-order to the second-order such that $(\alpha) \forall X \exists y (X \equiv f(y))$, where ‘ \equiv ’ expresses coextensiveness. Note that ‘ f ’ forms a predicate from a term, since it’s from the first-order to the second-order. First set $Cx =_{df} \neg f(x)x$. Pick any first-order y . Either Cy or $\neg Cy$. Consider the first case: Cy . By definition of C : $\neg f(y)y$. So $C \not\equiv f(y)$. Now consider the second case: $\neg Cy$. By definition of C : $\neg \neg f(y)y$. Hence $f(y)y$. So again $C \not\equiv f(y)$. Either way: $C \not\equiv f(y)$. Since y was arbitrary, we’ve got a counterexample to (α) : $\neg \exists y (C \equiv f(y))$. So f assigns no first-order representative to C . Since f was arbitrary, we can generalise: for every f assigning first-order representatives to second-order “entities”, some second-order C lacks first-order representative under f ; and since the first-order quantifiers in the argument are unrestricted, no f^* with more inclusive domain is available. So a first-order account of what second-order quantifiers range over is impossible, and the truth-conditions of second-order quantification cannot be explicated in first-order terms. It is also worth noting that, given the initial assumptions, this argument undermines all responses to the concept *horse* problem that permit singular terms with which to refer to—and hence objects corresponding to—each concept. Such proposals include (Wright, 1998, 85-90), (Hale and Wright, 2012), and (Hale, 2013, 31–34); though (Liebesman, 2015, §5) disagrees.

²³ Think of Basic Law V as the second-order principle: $\forall X, Y (\text{ext}(X) = \text{ext}(Y) \leftrightarrow X1 - 1Y)$. ‘ext’ is the ‘extension of’ functor, which forms a singular term from a predicate or predicate variable. ‘ $X1 - 1Y$ ’ says that the X s are in one-one correspondence with the Y s.

haps only first-order quantified sentences expresses existence claims,²⁴ and so second-order quantification doesn't really bring "ontological commitment", in Quine's (in)famous phrase. On this view, the **Criterion of Referentiality** is misapplied in the second-order case, and the apparently non-referential semantics really is non-referential.

Should we regard only first-order quantified sentences as expressing existence claims? I know of no compelling reason to do so.

The best reason I can find is that the English word 'exists' does not permit second-order usage. We can say that particular objects exist, e.g.: 'Al exists'. We can also say that objects of particular kinds exist, e.g.: 'cookies exist'. Both uses can be captured with first-order quantification, e.g.: ' $\exists x(x = \text{Al})$ ', ' $\exists x(\text{COOKIE}(x))$ '. But we cannot, in English, attribute existence to the predicational aspects of reality generalised by second-order sentences like ' $\exists X(X(\text{Al}))$ '. Strings like 'is a cookie exists' are not grammatical English. The point is not that second-order quantification isn't expressible in English. The point is that even if second-order quantification is expressible in English, it doesn't use the word 'exists'. So we speakers of English should not regard the contents of second-order quantified sentences as concerning existence.

The problem with this argument is its linguistic focus. It shows, at most, that the English word 'exists' can't be used to express second-order existential generalisation. It remains open that 'exists' expresses only the first-order restriction of the background fundamental notion: existential generalisation.²⁵ If so, it would be more perspicuous to identify existence with the more general and fundamental notion of which 'exists' expresses a restriction. On this view, the fully general notion of existence goes with existential quantification regardless of order, and regardless of expressibility with the English 'exists'.

The issue is liable to appear merely verbal. First- and second-order quantifiers both express kinds of generalisation. The question is whether they both express existence too. What turns on this issue? Why not regard both views as equally acceptable precisifications of our ordinary notion of existence? I see no reason not to do so. One precisification stays closer to ordinary usage, restricting 'existence' to the first-order. The other precisification expands ordinary usage to track the more general and fundamental underlying notion.²⁶ Both precisifications are equally legitimate. Dispute about which is correct is no more substantive than dispute about whether uses of 'mass' prior to the discovery of special relativity express relativistic mass or proper mass.²⁷ Prior usage didn't distinguish between two interpretations which we now see come apart. As a result, a semantic decision must now be made, with neither choice ruled out by earlier usage as incorrect. There

²⁴ More carefully: perhaps non-first-order quantified sentences expresses existence claims only if interpreted using first-order quantification. That's ruled out by the fundamental semantic distinction between first- and second-order quantification outlined at the beginning of this section.

²⁵ Although convenient, this probably isn't quite the right way to put it. One can analyse English 'exists'-sentences using first-order existential quantification without being committed to 'exists' expressing such quantification. It follows only that certain first-order quantified sentences express what English 'exists'-sentences express. The point in the text is then: even if no second-order quantified sentence expresses what any English 'exists'-sentence expresses, it remains open that English 'exists'-sentences express only the first-order restriction of the fundamental background notion of what's expressed by existentially generalised sentences regardless of order.

²⁶ §5.3 discusses a Fregean problem with this notion, and some potential solutions.

²⁷ See (Field, 1973) for discussion.

is thus a perfectly legitimate sense in which all orders of existential quantification express existence.²⁸ On this precisification, the **Criterion of Referentiality** classifies predicates governed by clauses like (Fii) and (Gii) as genuinely referential because those clauses entail the existence of second-order aspects of reality to which the predicates 2-refer; that's what's expressed by (Fii \exists) and (Gii \exists). Combining that with the view about second-order quantification from the beginning of this section, we obtain: two fundamentally different but intimately connected notions of existence, first- and second-order existence.

4 A Higher-Order Fregean Theory of Predication

I've argued that a seemingly non-referential semantics for predicates is really a referential semantics, one that sees the semantic roles of singular terms and predicates as involving parallel structures of word-world relations. That argument turned on views about the relationship between first- and second-order quantification, and between second-order existential quantification and existence, that I articulated in the preceding section. It is now time to return to the Fregean theory of predication. I'll use higher-order resources to formulate a version of that theory that survives refutation by the concept *horse* argument.²⁹

4.1 Objecthood, Concepthood, and Reference

This subsection looks at **Objecthood**, **Concepthood**, and **Reference**. The next turns to **Exclusivity**.

Our first task is to formulate a version of **Objecthood**. Some notation will be helpful. ' \diamond ' is a sentence operator expressing the "may in principle" modality in the original presentation of the theory. '1-REFERS' regiments the binary 1-reference predicate 'refers to', taking singular terms in both argument positions. 'OBJECT' regiments the monadic predicate 'is an object' taking singular terms in its argument position. We can now formulate **Objecthood** as:

Objecthood*: $\forall x(\text{OBJECT}(x) \leftrightarrow \diamond(\exists y : y \text{ is a singular term})(1\text{-REFERS}(y, x)))$.

Whatever a term may in principle 1-refer to is an object; nothing else is an object. Plausibly, a singular term can in principle be introduced for whatever's in the range of first-order quantification; nothing in the semantic role of singular terms precludes them from doing so, even if it's impossible for there to be language users that do introduce such terms. That is:

²⁸ More strongly: we should reject the ordinary notion of existence as lacking distinctive theoretical import because: (i) our ordinary word 'exists' can be precisified in two equally good ways; and (ii) both precisifications mark distinctions that are readily marked in other, less emotive and loaded, ways.

²⁹ As far as I'm aware, the closest extant proposal is due to Michael Dummett (1973, 211–218), which Dummett attributes to Frege. There isn't space here to properly compare our approaches. One key difference is that Dummett's goal is to explicate Frege's semantics and metaphysics within natural language. This brings commitments in natural language semantics—about the interpretations of English quantification and the copula *inter alia*—that I avoid. It is also unclear how to extend Dummett's proposal beyond the ordinary predicates considered so far to the rest of Frege's type-theoretic hierarchy. §5 discusses the corresponding generalisation of my proposal and addresses some recent criticisms of Dummett in (Hale and Wright, 2012, 96–98) and (Hale, 2013, 27–28).

- $\forall x \diamond (\exists y : y \text{ is a singular term})(1\text{-REFERS}(y, x))$.

That renders **Objecthood*** equivalent to:

Objecthood**: $\forall x(\text{OBJECT}(x))$.

First-order quantifiers range over objects.

Our second task is to formulate a version of **Concepthood**. To replace ‘is a concept’, which takes singular terms in its argument positions, I introduce a new theoretical expression ‘CONCEPT’ that takes a predicate in its sole argument position to form a sentence.³⁰ Recall that the appropriate reference predicate for predicates is ‘2-REFERS’. Thus **Concepthood** becomes:

Concepthood*: $\forall X(\text{CONCEPT}(X) \leftrightarrow \diamond(\exists y : y \text{ is a predicate})(2\text{-REFERS}(y, X)))$.

Whatever a predicate may in principle 2-refer to is a concept; nothing else is a concept.³¹ Plausibly, a predicate can in principle be introduced for whatever’s in the range of second-order quantification; nothing in the semantic role of predicates precludes them from doing so, even if it’s impossible for there to be language users that do introduce such predicates. That is:

- $\forall X \diamond (\exists y : y \text{ is a predicate})(2\text{-REFERS}(y, X))$.

That renders **Concepthood*** equivalent to:

Concepthood**: $\forall X(\text{CONCEPT}(X))$.

Second-order quantifiers range over concepts.

Here’s an objection. Because **Objecthood*** employs first-order quantification, it’s silent about what it takes for second-order aspects of reality to be objects. And because **Concepthood*** employs second-order quantification, it’s silent about what it takes for first-order aspects of reality to be concepts. In that sense, **Objecthood*** and **Concepthood*** are not fully general accounts of what it takes to be an object/concept. But, the objection continues, the intended force of the original **Objecthood** and **Concepthood** is fully general: they settle what it takes for any aspect of reality whatsoever to be an object/concept. To capture the intended force of **Objecthood** and **Concepthood**, we’d need to supplement **Objecthood*** and **Concepthood*** with:

³⁰ Might any natural language expression do the work of ‘CONCEPT’? Following Dummett (1973, 215), we might try: ‘for anything whatsoever, either it ... or it’s not the case that it ...’, where the dots mark a single argument position reserved for a predicate. Wright (1998, 78–79) observes that Dummett’s proposal’s adequacy depends on the Law of Excluded Middle (LEM). So he suggests a replacement: ‘either nothing ... or possibly something ...’. Just as Dummett’s proposal involves a non-trivial logical commitment, however, Wright’s involves a non-trivial metaphysical commitment: it precludes concepts that necessarily universally violate LEM; that is, concepts X such that, necessarily, for every object y , $X(y) \vee \neg X(y)$ fails (in whatever sense LEM was initially thought to fail and which motivated Wright’s proposal). It’s only a small step from countenancing concepts that violate LEM, to such necessary violators of LEM. An alternative that permits such concepts is ‘for any object whatsoever, it ... iff it ...’, reading ‘iff’ as the material biconditional. Yet even that is adequate only given the Law of Non-Contradiction; for each of its instances is equivalent to something of the form $\forall x \neg(F(x) \wedge \neg F(x))$. More generally, it is doubtful that any adequate analysis of ‘CONCEPT’ can or should entirely avoid non-trivial commitments of some kind.

³¹ Since the English ‘whatever’ is arguably first-order, this is only a rough gloss on **Concepthood***.

(O) $\forall X(\text{OBJECT}(X) \leftrightarrow \diamond(\exists y : y \text{ is a singular term})(1\text{-REFERS}(y, X)))$.

(C) $\forall x(\text{CONCEPT}(x) \leftrightarrow \diamond(\exists y : y \text{ is a predicate})(2\text{-REFERS}(y, x)))$

According to standard, and Fregean, views, however, neither principle is semantically evaluable. On such views, each argument position accepts expressions from exactly one semantic category; when argument positions are filled incorrectly, the result isn't semantically evaluable. Since (O) has a second-order variable in the argument positions of 'OBJECT' and '1-REFERS', it's not semantically evaluable. Since (C) has a first-order variable in the argument position of 'CONCEPT' and the second argument position of '2-REFERS', it's not semantically evaluable. Yet, the objection maintains, both principles are needed to capture the intended force of **Objecthood** and **Concepthood**, viz. fully general accounts of what it takes to be an object/concept.

The objection fails. To see why, suppose these failures of semantic evaluability arise from (i) merely grammatical limitations that rule the relevant strings non-sentences outside the scope of the compositional rules, or (ii) artificial restrictions on the compositional rules that prevent them from assigning truth-conditions in the relevant cases. Then semantic evaluability could be restored by adopting a more liberal syntax, or by lifting the restrictions on the compositional rules. The intended force of **Objecthood** and **Concepthood** could then be captured by conjoining them with (O) and (C) in the newly liberalised language. So if the objection is to stick, (O) and (C) must be unevaluable in principle, purely by virtue of the semantic roles of their constituent expressions. On this view, ordinary predicates take only singular terms in their argument positions because what they express—whatever that might be—is meaningfully attributable only to first-order aspects of reality. And if an expression takes only predicates in its argument position, that's because what it expresses—whatever that might be—is meaningfully attributable only to second-order aspects of reality. Contents attributing objecthood (concepthood) to non-first-order (non-second-order) aspects of reality do not exist. So no amount of tinkering with syntax or compositional rules will allow sentences to express them. That's why inserting an expression from the wrong semantic category into an argument position undermines semantic evaluability. Call this orthodox view *restrictivism about semantic evaluability*.

The argument position of 'OBJECT' requires a singular term (or first-order variable). And the argument position of 'CONCEPT' requires a predicate (or second-order variable). So by restrictivism: objecthood cannot be meaningfully predicated of reality's second-order aspects, and concepthood cannot be meaningfully predicated of its first-order aspects. There are no such contents. So **Objecthood*** and **Concepthood*** are not deficient in the way the objection requires. If objecthood cannot be meaningfully attributed to non-first-order aspects of reality, a fully general account of what it takes to be an object needn't settle its applicability to second-order aspects of reality; for there are no contents attributing objecthood to non-first-order aspects of reality to be ruled true or false. Likewise *mutatis mutandis* if concepthood cannot be meaningfully attributed to non-second-order aspects of reality. Given restrictivism about semantic evaluability, therefore, the objection fails: **Objecthood*** and **Concepthood*** supply fully general accounts of what it takes to be an object/concept.

We've just seen how restrictivism about semantic evaluability undermines the objection. What if restrictivism's false, and (some) argument positions can meaningfully accept expressions from multiple semantic categories? Although that's certainly non-standard, I know of no principled arguments against it. Some have recently even argued for it, e.g. (Magidor, 2009) and (Linnebo and Rayo, 2012, §4). Given this departure from Fregean orthodoxy, there's no obstacle to (O) and (C)'s semantic evaluability. So we can conjoin them with **Objecthood*** and **Concepthood*** to capture the intended force of the original **Objecthood** and **Concepthood**. Regardless of whether restrictivism holds, the objection therefore fails.

It remains only to formulate a higher-order version of **Reference**:

Reference*: $(\forall x : x \text{ is a predicate})(\exists Y : \text{CONCEPT}(Y))(2\text{-REFERS}(x, Y))$.

Given that **Concepthood*** is equivalent to **Concepthood****, that's equivalent to:

Reference**: $(\forall x : x \text{ is a predicate})\exists Y(2\text{-REFERS}(x, Y))$.

The first three principles of our higher-order Fregean theory of predication are now in place. Let us turn to **Exclusivity**.

4.2 Exclusivity

This subsection considers how to formulate a higher-order version of **Exclusivity**.

At a first pass, we might try:

Exclusivity*: $\neg\exists x(\text{OBJECT}(x) \wedge \text{CONCEPT}(x))$.

Exclusivity**: $\neg\exists X(\text{OBJECT}(X) \wedge \text{CONCEPT}(X))$.

Exclusivity* ensures that no first-order aspects of reality are object-concepts. **Exclusivity**** ensures that no second-order aspects of reality are object-concepts. Given restrictivism about semantic evaluability, however, neither principle is semantically evaluable; for **Exclusivity*** has a first-order variable in the argument position of 'CONCEPT', where a second-order variable is required; and **Exclusivity**** has a second-order variable in the argument position of 'OBJECT', where a first-order variable is required. One option is thus to be more liberal about semantic evaluability. But what if we follow Fregean orthodoxy and stick with restrictivism?

According to restrictivism, **Exclusivity*/**** are not semantically evaluable because there are no such contents. More generally: there is no such thesis as **Exclusivity** out there to express. The appearance otherwise results from eliding the distinctions between semantic categories at the heart of higher-order logic. That's all too easy if our semantic theorising is conducted in languages like English, where higher-order quantification is difficult and clumsy to express. 'Is a concept' is essential—certainly in practice, perhaps also in principle—for generalising about the semantics of predicates in such languages. We can, however, consistently accommodate the motivation for **Exclusivity** without admitting its existence. That motivation was rejection of all possible putative examples of object-concepts, in defence of **Linguistic Exclusivity**, given **Determination** (§1.2). Exactly what form should this rejection take?

Suppose I try to use a certain form of words to make an assertion. You want to reject my assertion. Here are two ways you can do so. You can reject my assertion as false. That involves recognising that I made an assertion with truth-evaluable content, and commits you to the truth of that content's negation. Alternatively, you can reject my form of words as failing to express a truth-evaluable content. Since you deny that my putative assertion has truth-evaluable content, you avoid commitment to the truth or falsity of any such asserted content or negation thereof.³²

Exclusivity was motivated by rejection of all possible putative examples of object-concepts. If this is rejection as false, it brings commitment to the negation of each such example, and hence also to the truth of **Exclusivity***/**. If, however, this is rejection as contentless, one avoids commitment to the existence of any contentful theses about object-concepts, whether presenting particular cases or generalising about them. One is instead committed to the non-existence of all such contents. There are, on this view, no truth-evaluable contents involving identities between objects and concepts. That includes **Exclusivity**. One can therefore respect the motivation for **Exclusivity** without being committed to its existence by rejecting as contentless all possible attempts to talk about things that are both objects and concepts.^{33, 34}

This approach departs from the initial Fregean theory of predication. That was a theory combining four propositions: **Reference**, **Objecthood**, **Concepthood**, and **Exclusivity**. Our new, higher-order approach combines only three propositions: **Reference***, **Objecthood***, and **Concepthood***. Advocates of the approach combine acceptance of those three propositions with rejection as contentless of all forms of words usable by their opponents to putatively talk about object-concepts. The truths about objects, concepts, and word-world relations can all be expressed in a higher-order language in which **Exclusivity** cannot be formulated.

To what extent is this Fregean? Without wanting to get bogged down in interpretative issues, this second version of **Exclusivity** does seem to accord with Frege's text.

“[W]hat is said here concerning a concept can never be said concerning an

³² (Rumfitt, 2014, §5) applies a similar notion of rejection to a Liar-like paradox.

³³ Talk about object-concepts used to explicate the view will have to be rejected as contentless too. So what kind of contents are we denying the existence of? §5 considers some related expressibility problems for higher-order Fregeans.

³⁴ Alongside rejection of **Exclusivity***/** as contentless, this approach also requires rejection as contentless of any other forms of words that could be used to talk about object-concepts. Notably, that includes a cross-order taking singular terms in one argument position and predicates in the other. In an interesting recent paper, Trueman (2014, §6) uses, in effect, restrictivism about semantic evaluability to argue against cross-order identity. His argument turns on the impossibility of formulating reflexivity for cross-order identity. It's impossible because it would require a single order of variable in each argument position of the cross-order identity predicate. Restrictivism then implies that there is no cross-order identity relation. I don't think that's decisive. Reflexivity involves a single quantifier binding two occurrences of the same variable, e.g.: $\forall x(x = x)$; $\forall X(X = X)$. Advocates of cross-order identity might reasonably take that to show that reflexivity is essentially intra-order. If so, then inability to formulate cross-order reflexivity is unproblematic for cross-order identity. More problematic, it seems to me, is the impossibility of formulating indistinguishability for cross-order identity, given restrictivism about semantic evaluability; for then ' $\forall\Phi(\Phi(X) \leftrightarrow \Phi(y))$ ' is not semantically evaluable, though a predicate expresses identity only if it guarantees complete indistinguishability.

object...I do not want to say it is false to say concerning an object what is said here concerning a concept; I want to say it is impossible, senseless, to do so.”
(Frege, 1892, 189)

Although Frege’s discussing one particular case, his point is clearly intended to hold generally. One thing that can meaningfully be said of an object is that it’s an object. So, on Frege’s view, one cannot meaningfully say of a concept that it’s an object. So one cannot meaningfully express **Exclusivity**. That’s not due to some deficiency of our language. It’s because no such thesis exists to be expressed.

4.3 The Concept *Horse* Argument Revisited

It’s now time to see how this higher-order Fregean theory of predication interacts with the concept *horse* argument. By way of reminder, here’s the argument:

- (1) Premiss: The predicate ‘est un cheval’ refers to the concept *horse*.
- (2) Premiss: ‘The concept *horse*’ is a singular term.
- (3) Premiss: ‘The concept *horse*’ refers to the concept *horse*.
- (4) By (1), **Concepthood**: The concept *horse* is a concept.
- (5) By (2), **Objecthood**: Whatever ‘the concept *horse*’ refers to is an object.
- (6) By (3), (5): The concept *horse* is an object.
- (7) By (6), **Exclusivity**: The concept *horse* is not a concept.
- (8) By (4), (7): The concept *horse* both is and is not a concept.

Higher-order Fregeans can resolve the problem by rejecting premiss (1).

Premiss (1) was motivated by **Reference**. If a predicate refers, then it refers to something. (1) was offered as a placeholder account of that something for ‘est un cheval’. What is (1)’s logical form? Since ‘the concept *horse*’ is a singular term, ‘refers’ in (1) must express 1-reference, otherwise it wouldn’t be well-formed. Where t is a singular term, (1) is thus of the form:

- 1-REFERS(‘est un cheval’, t).

Given our regimentation of **Reference** as **Reference***, however, that’s the wrong way to specify the referents of predicates. **Reference*** employs 2-reference not 1-reference. So the sense in which ‘est un cheval’ refers to something is not given by:

(1 \exists) $\exists x(1\text{-REFERS}(\text{‘est un cheval’}, x))$.

but by:

(1 \exists^*) $\exists X(2\text{-REFERS}(\text{‘est un cheval’}, X))$.

Since ‘est un cheval’ is a predicate, that follows from **Reference***. (1) was supposed to specify a referent for ‘est un cheval’ by giving an account of the something to which it refers. But (1) doesn’t instantiate $(1\exists^*)$; for (1) has 1-reference where it needs 2-reference, and consequently has a singular term where it needs a predicate. No principle of (1)’s form specifies a 2-referent for ‘est un cheval’. **Reference*** therefore provides no motivation for (1), or for any other such principle that attempts to use a singular term to specify a 1-or-2-referent for ‘est un cheval’. Higher-order Fregeans can thus escape the concept *horse* problem by rejecting both premiss (1) and $(1\exists)$ which motivated it.

Reference motivates (1) only if formulated in first-order terms as:

- $(\forall x : x \text{ is a predicate})(\exists y : y \text{ is a concept})(1\text{-REFERS}(x, y))$.

Since ‘est un cheval’ is a predicate, that does indeed imply $(1\exists)$. So the first-order formulation of **Reference** would motivate (1) as a specification of the something to which ‘est un cheval’ 1-refers. But higher-order Fregeans need have no truck with such first-order alternatives to **Reference***. As I argued in §2, expressions whose semantic clauses employ 2-reference are no less referential than those whose semantic clauses employ 1-reference: 2-referring to second-order aspects of reality is a genuine way of referring.

Higher-order Fregeans cannot simply leave it at that. They must also explicitly say to what ‘est un cheval’ refers. That’s easy to do with ‘2-REFERS’:

- 2-REFERS(‘est un cheval’, is a horse).

More long-windedly: $\forall x(\text{‘est un cheval’ applies to } x \leftrightarrow x \text{ is a horse})$.

To say which second-order aspect of reality a predicate 2-refers to just is to use a predicate to delineate the objects to which it applies. The concept *horse* problem appeared difficult to resolve because the English ‘refers to’ requires—or, perhaps better: is most naturally regimented as requiring—a singular term in its second argument position. That is, the English ‘refers to’ expresses 1-reference not 2-reference. Given higher-order quantification, §3’s connection between second-order quantification and existence, and the **Criterion of Referentiality**, however, the core notion of reference is clearly present in both cases. Semantic clauses formulated using 1-reference and 2-reference are both open to existential generalisation, and so entail the existence of (first- or second-order) aspects of reality to which the relevant expression is related by virtue of playing its particular semantic role. Two mistakes thus drive the concept *horse* problem:

- Assimilating referentiality to one particular word-world relation, expressed by ‘1-refers’, rather than maintaining the connection between semantic role and word-world relations in general which the **Criterion of Referentiality** captures.
- Failing to recognise the existential import of second-order quantification and its connection with irreducibly second-order existence.

Having rectified those mistakes, we should see nothing amiss in using predicates to specify the (2-)referents of predicates, just as we use singular terms to specify the (1-)referents of terms. ‘The concept *horse*’, ‘is a concept’, and ‘(1-)refers to’ play no role in the resulting referential semantics for predicates. The concept *horse* problem therefore does not afflict this approach.

5 Expressibility Problems

My higher-order Fregean proposal is now complete. To close the paper, I consider a prominent style of objection to views of broadly this kind. §5.1 outlines some versions of the problem that occur in the literature, to which §5.2 responds. §5.3 then turns to a more difficult version of the problem and some speculative remarks about how to resolve it.

5.1 The Problem

The problem is that certain truths about the higher-order Fregean framework are, by its own lights, in principle inexpressible. Moreover, the objection continues, these truths are essential to the framework's intended significance. In that sense, the framework appears self-undermining.

Our focus thus far has been on ordinary, or *first-level*, predicates whose argument positions take singular terms. The expressibility problem arises from considering other predicate-like expressions. *Second-level* predicates take first-level predicates in their argument positions to form sentences. *Third-level* predicates take second-level predicates in their argument positions to form sentences. In general, *nth-level* predicates take $(n - 1)^{th}$ -level predicates in their argument positions to form sentences. These levels are also known as *types*. In a language, like our higher-order metalanguage, where syntax mirrors semantics, these types amount to semantic categories. Note that second-order quantification is quantification into first-level predicate position. Similarly, *third-order* quantification is quantification into second-level predicate position, and *nth-order* quantification is quantification into $(n - 1)^{th}$ -level predicate position.

On Frege's view, this linguistic hierarchy of categories of expression runs parallel to an ontological hierarchy of categories of entity. We've been focusing on objects and *first-level* concepts applicable to objects. *Second-level* concepts are applicable to first-level concepts. *Third-level* concepts are applicable to second-level concepts. In general *nth-level* concepts are applicable to $(n - 1)^{th}$ -level concepts. The *nth-level* concepts are exactly the potential referents for *nth-level* predicates.³⁵ Within each hierarchy, no two levels overlap.

The expressibility problem is this: by the Fregean framework's own lights, certain core structural features, just outlined, of these Fregean hierarchies and their inter-relations cannot be expressed. Putative examples from the literature include:³⁶

- (A) First-level predicates refer to (first-level) concepts.
- (B) First-level concepts are had only by objects.
- (C) Second-level concepts are had only by first-level concepts.

³⁵ This will need complicating to accommodate cross-level expressions like '2-REFERS' and the cross-level relations to which they apply. See (Williamson, 2013, 221–222) for a suitable approach, drawn from (Gallin, 1975, 67–78).

³⁶ (A)–(C) are from (Hale and Wright, 2012, 97, 98, 104 respectively). (A) and (B) are from (Hale, 2013, 27–28). (D)–(F) are from (Linnebo, 2006, §6.4). I've changed terminology and wording to fit present usage.

- (D) The referent of a complex expression is a function of the referents of its component expressions.
- (E) Each expression of each semantic type has a unique referent, unique both within that entity's ontological category and from across all categories.
- (F) There are infinitely many ontological categories of potential referents, one for each semantic type of expression, with different categories supplying potential referents for expressions of different types (as the objects are the potential referents of singular terms, the first-level concepts are the potential referents of first-level predicates, the second-level concepts are the potential referents of second-level predicates, etc.).

I'll now argue that (A)–(C) are expressible under my proposal. The difficult cases are really (D)–(F).

First, some notation. Superscripts indicate levels, e.g.: ' X^1 ' is a first-level predicate variable; ' X^2 ' is a second-level predicate variable. Term-position variables remain in lower case without superscript, though they can also be treated as level 0. I won't use superscripts on cross-level predicates like '2-REFERS', and ignore quantification into cross-level positions; I want to avoid complicating the hierarchy and notation more than is absolutely required, and cross-level relations raise no new problems of principle.

§4 showed how to accommodate (A) as **Reference***. We also need to accommodate principles like:

- (Aii) Second-level predicates refer to second-level concepts.
- (Aiii) Third-level predicates refer to third-level concepts.

Here's how to do so.

First, we need higher-level counterparts of 'is a second-level concept' and 'is a third-level concept'. §4.1 introduced the monadic second-level predicate 'CONCEPT' to replace the first-level 'is a concept'; I henceforth write it as ' 1-CONCEPT^2 '. We can also introduce similar higher level predicates for higher levels of concept:

- ' 2-CONCEPT^3 ': monadic third-level predicate whose argument position accepts second-level predicates; replaces 'is a second-level concept'.
- ' 3-CONCEPT^4 ': monadic fourth-level predicate whose argument position accepts third-level predicates; replaces 'is a third-level concept'.

Just as ' 1-CONCEPT^2 ' satisfies ' $\forall X^1(1\text{-CONCEPT}^2(X^1))$ ', these are governed by:

- $\forall X^2(2\text{-CONCEPT}^3(X^2))$.
- $\forall X^3(3\text{-CONCEPT}^4(X^3))$.

And so on upwards.

Next, we need a reference predicate for second-level predicates. We can obtain one using the same technique as in §2 for the 2-reference predicate '2-REFERS' appropriate for

first-level predicates. Formulate the semantic clauses for second-level predicates P^2 so that they combine with (i) the semantic clauses for first-level predicates P^1 , and (ii) the compositional principle for predications $\ulcorner P^2(P^1) \urcorner$, to yield (iii) truth-conditions for such predications directly, without additional semantic machinery (such as the account of applicability needed to supplement (Fi) but not (Fii)). Just as semantic clause (Fii) for the first-level predicate ‘ F ’ in §2 used a first-level predicate applicable to objects, these semantic clauses for second-level predicates will employ second-level predicates applicable to 1-concepts. Let A be such a clause. Replace the second-level predicate in A with a new second-level predicate variable. Replace the singular term for the second-level predicate with a new first-order variable. The result is an open sentence $\ulcorner A(x, Y^2) \urcorner$ with ‘ x ’ and ‘ Y^2 ’ free. This is (alternatively: defines) the reference predicate for second-level predicates, which I henceforth write as ‘3-REFERS(x, Y^2)’. It expresses a cross-level relation of 3-reference between expressions x and 2-concepts Y^2 . This technique extends naturally to all levels.³⁷

We can now capture (Aii) and (Aiii) as:

(Aii*) $(\forall x : x \text{ is a second-level predicate})(\exists Y^2 : 2\text{-CONCEPT}^3(Y^2))3\text{-REFERS}(x, Y^2)$.

(Aiii*) $(\forall x : x \text{ is a third-level predicate})(\exists Y^3 : 3\text{-CONCEPT}^4(Y^3))4\text{-REFERS}(x, Y^3)$.

We can also capture (B) and (C) as:³⁸

(B*) $(\forall X^1 : 1\text{-CONCEPT}^2(X^1))\forall y(X^1(y) \rightarrow \text{OBJECT}^1(y))$

(C*) $(\forall X^2 : 2\text{-CONCEPT}^3(X^2))\forall Y^1(X^2(Y^1) \rightarrow 1\text{-CONCEPT}^2(Y^1))$

Similarly for all levels.³⁹

³⁷ Here’s a concrete example. To aid readability, I mix formal and informal vocabulary, and take some grammatical liberties. Compositional principle for second-level atomic predications $\ulcorner P^2(P^1) \urcorner$: $\ulcorner P^2(P^1) \urcorner$ is true iff P^2 2-applies to the 2-referent of P^1 . ‘2-applies’ is the higher-level (and cross-level) analogue of the familiar ‘applies to’; its second argument position is for first-level predicates. Semantic clause for monadic ‘ F^1 ’: ‘ F^1 ’ 2-refers to is a horse. Semantic clause for monadic ‘ G^2 ’: $\forall Z^1('G^2'$ 2-applies to $Z^1 \leftrightarrow H^2(Z^1)$); that is, ‘ G^2 ’ 2-applies to every 1-concept that H^2 s (where ‘ H^2 ’ is an interpreted second-level predicate of the metalanguage, e.g. ‘something’). These semantic clauses and compositional principle together assign the following truth-condition to ‘ $G^2(F^1)$ ’ without the need for any additional semantic principles such as an analysis of 2-application: ‘ $G^2(F^1)$ ’ is true iff ‘ G^2 ’ 2-applies to the 1-concept that ‘ F^1 ’ 2-refers to; iff H^2 (is a horse). The reference predicate ‘3-REFERS(x, Y^2)’ for second-level predicates is defined by replacing ‘‘ G^2 ’ with ‘ x ’ and ‘ H^2 ’ with ‘ Y^2 ’ in the semantic clause for ‘ G^2 ’ to obtain ‘ $\forall Z^1(x$ 2-applies to $Z^1 \leftrightarrow Y^2(Z^1)$ ’.

³⁸ Hale and Wright (2012, 104) use not (B) but ‘first-level functions can take only objects as arguments’. Since the functions are a subclass of the relations, this amounts to a claim about relational 1-concepts. And since it’s couched in terms of what these functions take as arguments, rather than what they map to the True, Hale & Wright’s claim is slightly more general than (B*). Assuming classical logic, a better regimentation of (a monadic version of) Hale & Wright’s principle is thus: $(\forall X^1 : 1\text{-CONCEPT}^2(X^1))\forall y((X^1(y) \vee \neg X^1(y)) \rightarrow \text{OBJECT}^1(y))$. I ignore this complication in the main text.

³⁹ Objection: (B*) doesn’t capture (B)’s intended force; because it employs first-order quantification over objects, it doesn’t prevent 1-CONCEPTS from holding of non-objects that aren’t first-order aspects of reality, whereas (B) does. Reply: this goes the same way as for (O) and (C) in §4.1. If restrictivism about semantic evaluability holds, there are no contents attributing 1-concepts to non-first-order aspects of reality which need ruling out as false. If restrictivism fails, we can supplement (B*) with principles like the following to capture (B)’s intended force: $(\forall X^1 : 1\text{-CONCEPT}^2(X^1))\forall Y^1(X^1(Y^1) \rightarrow \text{OBJECT}^1(Y^1))$. Either way, the objection fails. Similarly for the parallel objection to (C*).

The really problematic cases are (D)–(F). They’re problematic because higher-order logic is a form of type theory, and (D)–(F) essentially involve two features disallowed by standard formulations of type theory. First feature: cross-type quantification over potential referents for expressions from multiple semantic categories; quantification is only ever over one ontological category. Second feature: cross-type predication of a single property or relation—e.g. reference, being a function of—of potential referents for expressions from multiple semantic categories; no relation of any category takes more than one category of entity in any of its argument positions. These are disallowed because each argument position accepts exactly one semantic category of expression. These type-restrictions would have to be violated by sentences expressing (D)–(F). Given restrictivism about semantic evaluability, it follows that there are no such contents as (D)–(F).⁴⁰ There isn’t even a univocal notion of entity or thing, applicable within each ontological category; there are really a multiplicity of notions, corresponding to the multiplicity of orders of quantification into different semantic categories. So the problem isn’t primarily about linguistic inexpressibility; it’s about non-existence of the relevant contents to be expressed.

5.2 Living with Inexpressibility

One option is to reject the restrictivism about semantic evaluability on which the argument from inexpressibility to non-existence depends. That’s out of line with orthodox Fregeanism, however, and the resulting logico-metaphysical picture quickly becomes complicated. So I won’t consider it here. I instead consider the following response: accept the non-existence of contents (D)–(F), and deny that adequate higher-order Fregeanism requires them. Those principles capture core structural features of the twin Fregean ontological and linguistic hierarchies, and their inter-relations. So to make the response stick, we need to show how to capture that structure in terms acceptable to higher-order Fregeans, hence without (D)–(F). This subsection shows how to do so. The next turns to a more difficult case.

Although higher-order Fregeans cannot capture exactly the content of (D)–(F), they can express type-restricted replacements. For example, the following type-restricted replacements for (E) and (F) are expressible in our setting:

(E*) Each singular term 1-refers to a unique object, but doesn’t 2-refer to any 1-concept, or 3-refer to any 2-concept, or⁴¹

Each first-level predicate 2-refers to a unique 1-concept, but doesn’t 1-refer to any object, or 3-refer to any 2-concept, or

⁴⁰ The same goes for these generalisations of (A)–(C): $\forall i(i\text{-level predicates } (i + 1)\text{-refer to } i\text{-level concepts})$; $\forall i(i\text{-level concepts are had only by } (i - 1)\text{-level concepts})$. I don’t discuss these explicitly because they raise no new issues beyond those of (D)–(F).

⁴¹ Note that since the 1-reference predicate requires a singular term (or first-order variable) in its second argument position, no contents predicate, of any singular term, 1-reference to a concept (of any level). So no such contents need ruling out as false to capture the intended import of (E). Likewise for 2-reference to things other than 1-concepts, 3-reference to things other than 2-concepts, etc. See §4.1’s discussion of (O) and (C) for more.

Each second-level predicate 3-refers to a unique 2-concept, but doesn't 1-refer to any object, or 2-refer to any 1-concept, or

(F*) There are infinitely many different semantic types of expression: singular terms, first-level predicates, second-level predicates,

Reality exhibits a parallel structure, in the following sense: there are objects (which are potential 1-referents of all and only singular terms), and there are 1-concepts (which are potential 2-referents of all and only first-level predicates), and there are 2-concepts (which are potential 3-referents of all and only second-level predicates), and there are

Note that (E*) and (F*) are infinitely long. So finite beings like ourselves cannot utter them. Is that problematic? No. This inexpressibility is consistent with the existence of the relevant contents; it threatens only their expressibility by limited beings like ourselves, using certain finite resources. The heart of the problem isn't about expressibility *per se*, or *our* ability to linguistically capture the Fregean hierarchies. It's about what truths and contents about those hierarchies exist; for our interest here is in the metaphysics of language, and there is no good reason to expect that to be restricted by what beings like us can say. Failure of finite expressibility does not entail that reality lacks the relevant metaphysical structure.

What about (D)? It amounts to the following claim—or collection of claims, one for each value of '*n*'—about the relationship between the referents of complex expressions and those of their components:

(D*) There is a relation R such that:

- For any x_1, \dots, x_n, y, z : if $R(x_1, \dots, x_n, y)$ and $R(x_1, \dots, x_n, z)$, then $y = z$.
- For any complex expression e of any type with components e_1, \dots, e_n , and for any x_1, \dots, x_n : if e_1 refers to x_1, \dots , and e_n refers to x_n , then e refers to any y such that $R(x_1, \dots, x_n, y)$.

The first bullet says that R is a function. The second says that R determines the referents of complexes from those of their components. Since e, e_1, \dots, e_n may be of any types, the other quantifiers (including the R -quantifier) must range over, and the relevant reference relation must be able to hold of, potential referents for expressions of every type.

Higher-order Fregeans can avoid these commitments of (D) by following the same strategy as for (E) and (F). Instead of one principle covering all types, use a series of type-restricted principles, one for each way of forming a complex expression from simpler ones. These type-restricted principles will employ different reference predicates and orders of quantifier, depending on the types involved. I won't spell them out in detail here, since they raise no new problems of principle.

Is there any reason to prefer (D)–(F) over these type-restricted alternatives? Those replacements amount to infinite sequences of principles, different principles for different types. Cross-type quantification provides a way to combine the members of each sequence into a single quantificational claim, i.e. (D)–(F). Doing so reveals commonality between

the semantic roles of different types that is invisible to opponents of cross-type quantification. This suggests two reasons to prefer (D)–(F).

First reason: although expressions of different types play different semantic roles, there is nonetheless commonality amongst them. Yet such semantic commonality is exactly what higher-order Fregeans reject.⁴² And it goes hand-in-hand with §3's conception of higher orders of quantification as irreducible to lower orders. Absent independent argument for such commonality, higher-order Fregeans should be unmoved by this first reason.

Second reason: a semantic theory with (D)–(F) exhibits the theoretical virtue of simplicity to a greater degree than would one with only their type-restricted replacements. The strength of this consideration depends on the degree of increased simplicity, and how greatly one values it. Although the increase in simplicity is real, it's not so significant that its value in this particular case is that great; though that's perhaps a matter of personal theoretical taste. Moreover, the virtue of increased simplicity must be balanced against the vices of alternative solutions to the concept *horse* problem. Unless some solution is absolutely free from unattractive features, the disunity of these replacements for (D)–(F) is hardly decisive against higher-order Fregeans.

5.3 Expressing the Criterion of Referentiality

The **Criterion of Referentiality** presents a more difficult problem. Recall:

Criterion of Referentiality: An expression is referential iff non-trivial existential generalisation into its semantic clause is possible.

The **Criterion** was central to §2's argument that '2-REFERS' is a genuine reference predicate: it expresses a relation between expressions and reality's second-order aspects, as shown by the entailment from '2-REFERS('est un cheval', is a horse)' to ' $\exists X^1$ 2-REFERS('est un cheval', X^1)'. The idea was to connect referentiality with the role of word-world relations in semantic role, rather than with one particular word-world relation drawn from one particular level in the Fregean ontological hierarchy. Such relations are revealed by non-trivial existential generalisation into semantic clauses. This argument is now under threat.

The **Criterion** is problematic because the present setting has no space for a fully general notion of existential generalisation applicable to all orders. First-order existential quantifiers take first-level predicates in their argument positions; so they are second-level predicates that 3-refer to 2-concepts. Second-order existential quantifiers take second-level predicates in their argument positions; so they are third-level predicates that 4-refer to 3-concepts. Since no two levels overlap, no notion of quantification covers both the first- and second-order.

Which notion of quantification appears in the **Criterion**? No answer classifies more than one semantic category as referential. So either the **Criterion** fails to provide a definite, unambiguous account of referentiality, or singular terms and predicates don't both refer in the intended sense. Replacing the **Criterion** with a series of type-restricted principles, as

⁴² Exception: each type's semantic role involves some word-world relation, in line with the **Criterion of Referentiality**. §5.3 discusses this issue.

in the previous subsection, would leave us without a univocal sense in which expressions of different types count as referential.

Likewise for the notion of word-world relation used to motivate the **Criterion**. There are first-level relations: potential 2-referents of first-level polyadic predicates, applicable to objects, and over which second-order quantifiers range. There are also second-level relations: potential 3-referents of second-level polyadic predicates, applicable to first-level concepts and relations, and over which third-order quantifiers range. And so on upwards. Since no two ontological levels overlap, no notion of relation covers multiple levels. So no notion of word-world relation covers the semantic roles of multiple categories of expression.

I see two kinds of potential solution to this problem.

The first solution attempts to find commonalities amongst different orders of existential quantifier, to justify using the same label ‘existential quantifier’ in each case. These commonalities cannot amount to a single classification or kind under which all existential quantifiers fall; for that would involve a single concept possessed by them all, contrary to the Fregean ontological picture. Different quantifier-concepts may, however, exhibit strictly different but nonetheless analogous features. These analogies may then suffice for grouping them naturally together.

As far as I can see, the most promising account of such commonality appeals to similarity of inferential role.⁴³ Let A_1 be a collection of argument forms featuring expression e_1 . Let A_2 be a collection of argument forms featuring expression e_2 . Now suppose that: all arguments of a form in A_1 are valid iff e_1 3-refers to a 2-concept that is a first-order existential quantifier; all arguments of a form in A_2 are valid iff e_2 4-refers to a 3-concept that is a second-order existential quantifier. Finally, suppose each of A_1 and A_2 is an instance of a single pattern. Then 3-reference to a first-order existential quantifier validates the same kind of arguments as does 4-reference to a second-order existential quantifier. Different semantic relationships to different levels of quantifier validate the same kinds of arguments. We don’t have a single classification applicable to each case because the relevant quantifiers and semantic relations come from different levels. But the valid argument forms for each order of quantifier-expression are of the same kind. In that sense, each of these orders of quantifier exhibits analogous behaviour. If the approach can be extended to all orders, we will have a sense in which they all count as existential quantifiers: they have analogous logical properties, which manifest in their validating the same kinds of argument.⁴⁴

That first approach attempts to vindicate the **Criterion of Referentiality** by finding a univocal sense in which higher-order Fregeans can regard different orders of existential quantifiers as existential quantifiers. An alternative strategy bypasses the **Criterion**, focusing instead on other commonalities amongst the ‘*i*-REFERS’ predicates. One promising candidate is this: when our metalanguage contains the object-language, each ‘*i*-REFERS’ predicate occurs in true disquotational principles for expressions from the category for

⁴³ The proposal draws on (Wright, 2007, §2).

⁴⁴ In the existential case, the relevant kind of argument forms will plausibly be: (i) from an instance to its existential generalisation; and (ii) from an existential generalisation to whatever follows from a suitably arbitrary instance.

which it's intended to serve as a reference predicate.⁴⁵ For example:

- 1-REFERS('Al', Al).
- 2-REFERS('eats cookies', eats cookies).

More long-windedly: $\forall x$ ('eats cookies' applies to $x \leftrightarrow x$ eats cookies).

The idea is that this commonality in the disquotational behaviour of the '*i*-REFERS' predicates suffices for regarding each as a genuine reference predicate.

Two responses to the problem of expressing the **Criterion of Referentiality** are now in place. Although I've presented them separately, they're really complementary. One respect of commonality amongst the '*i*-REFERS' predicates is that each admits existential generalisation into its second argument position. That's why *i*-referring expressions satisfy the **Criterion**. Making sense of this commonality amongst '*i*-REFERS' predicates requires a notion of existential quantifier that covers all orders. The two responses simply emphasise different aspects of reference's theoretical role, seeking to show how it's satisfied by each *i*-reference relation.

This subsection has been somewhat programmatic, leaving many interesting and important questions unanswered. How exactly should the valid argument forms for quantifiers be characterised so as to cover all orders? What is the underlying connection between disquotation and reference, such that disquotational commonality suffices for referential commonality? Do the analogies between different levels of quantifiers and reference predicates really warrant use of the same label 'quantifier' or 'reference predicate' in each case? Or is that just double-think, possible only by violating the Fregean's own strictures on what there is to be said? Whether the concept *horse* problem undermines the Fregean conception of language, reality, and their inter-relations turns on the availability of suitable answers to these questions. I cannot hope to settle that here.

References

- Beaney, M., editor (1997). *The Frege Reader*. Blackwell Publishers.
- Dummett, M. (1973). *Frege: Philosophy of Language*. Duckworth. Second (1981) edition.
- Field, H. (1973). Theory change and the indeterminacy of reference. *The Journal of Philosophy*, 70(14):462–481.
- Frege, G. (1892). On concept and object. *Vierteljahrsschrift für Wissenschaftliche Philosophie*, 16:192–205. Page references are to the version in (Beaney, 1997, 181–193).
- Gallin, D. (1975). *Intensional and Higher-Order Modal Logic: With Applications to Montague Semantics*. North-Holland.

⁴⁵ When metalanguage doesn't contain object-language, we can formulate similar true principles in which the second argument-position of '*i*-REFERS' is occupied by a translation into the metalanguage of the expression named by the occupant of its first argument-position. For example: 2-REFERS('est un cheval', is a horse). For further development of the connection between reference and disquotation, see (Krämer, 2013) and (Trueman, MSa).

- Hale, B. (2013). *Necessary Beings: An Essay on Ontology, Modality, and the Relations Between Them*. OUP.
- Hale, B. and Wright, C. (2001). *The Reason's Proper Study*. OUP.
- Hale, B. and Wright, C. (2012). Horse sense. *The Journal of Philosophy*, 109(1/2):85–131.
- Krämer, S. (2013). Semantic values in higher-order semantics. *Philosophical Studies*, 168(3):709–724.
- Liebman, D. (2015). Predication as ascription. *Mind*, 124(494):517–569.
- Linnebo, Ø. (2006). Sets, properties, and unrestricted quantification. In Rayo, A. and Uzquiano, G., editors, *Absolute Generality*, chapter 6. OUP.
- Linnebo, Ø. and Rayo, A. (2012). Hierarchies ontological and ideological. *Mind*, 121(482):269–308.
- MacBride, F. (2008). Predicate reference. In Smith, B. C. and LePore, E., editors, *The Oxford Handbook of Philosophy of Language*, chapter 19. OUP.
- MacBride, F. (2011). Impure reference: a way around the concept *Horse* paradox. *Philosophical Perspectives*, 25 (Metaphysics)(1):297–312.
- Magidor, O. (2009). The last dogma of type confusions. *Proceedings of the Aristotelian Society*, 109(1):1–29.
- Oliver, A. (2005). The reference principle. *Analysis*, 67:177–187.
- Prior, A. N. (1971). *Objects of Thought*. OUP.
- Rayo, A. and Williamson, T. (2003). A completeness theorem for unrestricted first-order languages. In Beall, J. C., editor, *Liars and Heaps: New Essays on Paradox*, chapter 15. OUP.
- Rayo, A. and Yablo, S. (2001). Nominalism through de-nominalization. *Nous*, 35(1):74–92.
- Rumfitt, I. (2014). Truth and meaning. *Proceedings of the Aristotelian Society: Supplementary Volume*, 88(1):21–55.
- Schiffer, S. (2003). *The Things We Mean*. OUP.
- Shapiro, S. (1991). *Foundations Without Foundationalism: a Case for Second-Order Logic*. OUP.
- Trueman, R. (2014). The concept horse with no name. *Philosophical Studies*, 172(7):1889–1906.
- Trueman, R. (MSa). Realism and the truth in nominalism.

- Trueman, R. (MSb). Substitution in a sense.
- Williamson, T. (2003). Everything. *Philosophical Perspectives*, 17(1):415–65.
- Williamson, T. (2013). *Modal Logic as Metaphysics*. OUP.
- Wright, C. (1998). Why Frege did not deserve his *Granum Salis*: a note on the paradox of ‘the concept horse’ and the ascription of *Bedeutungen* to predicates. *Grazer Philosophische Studien*, 55:239–263. Reprinted in (Hale and Wright, 2001, ch3).
- Wright, C. (2007). On quantifying into predicate position. In Leng, M., Paseau, A., and Potter, M., editors, *Mathematical Knowledge*, pages 150–174. OUP.