

*Penultimate draft—please cite/quote from the published version.*

RULE-FOLLOWING AND METAONTOLOGY  
T. Parent

At minimum, Wittgenstein's rule-following argument suggests that linguistic understanding does not consist in knowing interpretations. Kripke adapted the argument to say that linguistic meaning cannot be metaphysically fixed by interpretations.<sup>1</sup> Here, a different application of the argument will be attempted—the rule-following reasoning will be used to suggest that certain metaphysical questions cannot be answered by interpretations. It is my aim to present this version of the argument, and to say something useful about it.

To avoid quick dismissals, I wish to state that I am not a Wittgensteinian, nor a Kripkensteinian. Nevertheless, rule following seems to generate a noteworthy problem for metametaphysics. I ultimately reject the regress that looms, yet a “skeptical solution” will eventually be conceded.

The paper can seem tendentious in another respect. A striking fact about the contemporary scene is the popularity of metaphysics. The upshot of the present discussion, however, could be regarded as anti-metaphysical. But I am not hostile to everything that might be called “metaphysics.” The paper should instead be seen as a plea for clarity on what metaphysics can accomplish exactly.

---

<sup>1</sup> L. Wittgenstein, *Philosophical Investigations*, ed. G.E.M. Anscombe, P.M.S. Hacker, and Joachim Schulte, 4th ed. (West Sussex: Wiley-Blackwell, 1953). S. Kripke, *Wittgenstein on Rules and Private Language* (Cambridge, MA: Harvard University Press, 1982).

## 1. Preliminaries

To orient ourselves, consider: Can semantics serve the aims of metaphysics? The orthodox answer is, to a large extent, “yes.” David Lewis, for instance, argues that possible worlds are indispensable to modal semantics, and that their semantic utility amounts to a metaphysical argument for their existence.<sup>2</sup> In a related vein, meta-ethicists and epistemologists often trade the question of what normative facts are with a question of what normative terms mean, under the assumption that to answer the latter is (more or less) to answer the former.

That semantics can serve metaphysics is also assumed in ontology by both Quine and Carnap.<sup>3</sup> In Quine, the semantics of the quantifier determine the ontological commitments of a theory—but similarly, Carnap determines ontological commitments relative to a “framework,” a set of linguistic rules for the proper use of expressions. Such rules determine that, e.g., in the mathematics framework, ‘There is an even prime’ is analytic. In addition, Carnap jettisons “external” ontological questions for semantic reasons. Outside a framework, there are no rules for the use of expressions—so here, the sentences come out as “meaningless.”

Semantics in the service of metaphysics has been overtly embraced recently by Sider.<sup>4</sup> A primary thesis of his is that “every non-fundamental truth holds in virtue of some fundamental truth,” where, a fundamental truth is a “metaphysically basic or rock-bottom truth.”<sup>5</sup> The very idea of a “metaphysically basic truth” already shows some cross-over between metaphysics and semantics. But the co-mingling is even more explicit when Sider declares that, in defending his

---

<sup>2</sup> *On the Plurality of Worlds* (New York: Basil Blackwell, 1986).

<sup>3</sup> W. Quine “On What There Is,” *Review of Metaphysics* 2 (1948): 21–38. R. Carnap, “Empiricism, Semantics, and Ontology,” *Revue Internationale de Philosophie* 4 (1950): 20–40.

<sup>4</sup> *Writing the Book of the World* (New York: Oxford University Press, 2011).

<sup>5</sup> *op. cit.*, 112.

thesis, he does not seek metaphysical reductions between facts of various types. Rather his approach is:

that of a *metaphysical semantics*—a certain sort of semantic theory...with two distinctive features. First, meanings are to be given in purely joint-carving terms... Second...by giving meanings in fundamental terms it seeks...to show how what we say fits into fundamental reality.<sup>6</sup>

Sider's notion of "fit" may be a bit unclear—but presumably, metaphysical semantics should at least answer questions about the basic or fundamental ontology of the language. We could ask: Is the ontology of the language that of mereological nihilism? Of Berkeleyan idealism? Of Pythagoreanism? Or what? Sider himself eventually favors a kind of nihilism,<sup>7</sup> and presumably it is his metaphysical semantics that presents this answer.

At any rate, it is an important case where semantics is thought serviceable to metaphysics.<sup>8</sup> Further examples abound.<sup>9</sup> Yet my concern is not with all semantic theorizing that

---

<sup>6</sup> Ibid.

<sup>7</sup> op. cit., ch. 13.

<sup>8</sup> Sider makes clear, however, that he thinks the prospects for his metaphysical semantics are modest. But his reason seems to be that a metaphysical semantics will (of practical necessity?) be limited to "toy" examples. He does not seem to notice the potential for a rule-following problem. It is true he addresses the potential for "infinite descent," but the issue here is importantly different, as I shall explain in section 4.

<sup>9</sup> E.g., Davidson "The Method of Truth in Metaphysics," in *Inquiries into Truth and Interpretation* (Oxford: Clarendon Press, 1977), 199–214.: "[I]n making manifest the large features of our language, we make manifest the large features of reality. One way of pursuing metaphysics is therefore to study the general structure of our language" (p. 199). Also, Ludlow *Semantics, Tense, and Time: An Essay in the Metaphysics of Natural Language* (Cambridge, MA: MIT Press, 1999): "[C]oncrete questions about the nature of reality can be illuminated by what we know about semantic theory" (p. 5).

metaphysicians might engage in. The goal is narrower; it is just to illustrate that, *if* the rule-following problem is real, semantics could not serve metaphysics in at least one respect. Specifically, the reasoning suggests that seeking answers about fundamental ontology, in the form of semantic questions, leads to a vicious regress.

In Wittgenstein, the rule following argument works against the “interpretation theory” of linguistic understanding, where understanding the use of an expression consists in knowing its interpretation. The problem is that, on this view, speakers must know infinitely many expressions to know even one expression. Briefly, that’s because an interpretation is itself an expression, and thus would require knowledge of a second interpretation, which would require a third, and so on. In order to draw out parallels later, we may reconstruct the argument as follows.<sup>10</sup>

(1) Understanding an expression requires understanding the rule for its use. [Assume]

(2) Understanding the rule for an expression requires understanding an interpretation of the expression. [Assume for reductio]

(2.1) Then, understanding a particular expression  $E$  requires understanding its interpretation  $S$ . [From (1), (2)]

(2.2) And understanding  $S$  requires understanding an interpretation of  $S$ , say  $S_I$ . [From (1), (2)]

(2.3) And understanding  $S_I$  requires understanding its interpretation of  $S_I$ , call it  $S_2$ . [From (1), (2)]

(Etc...)

---

<sup>10</sup> The reconstruction is not intended to be serious Wittgenstein exegesis, though I do not think it is too much of a distortion. But my primary aim is just to clarify what I mean by “the rule-following argument.” But, n.b., sometimes I speak of the “rule following *problem*,” which refers just to the regress that is generated. The final statement at (4) (or at (4\*) below) I take as the *solution* to the problem.

(3) So, understanding *E* requires understanding infinitely many interpretations.

[From the preceding]

(4) So there is a way of understanding a rule that does *not* require understanding an interpretation.

[By reductio, (2)-(3)]<sup>11</sup>

Suppose (if only for the moment) that the argument shows that linguistic understanding is not just a matter of knowing interpretations. Then, one reaction might be: “ok, but what does the argument show for knowledge that patently is knowledge of interpretations?”

Where  $\Phi$  and  $\Psi$  are terms, let an “interpretation” be a true sentence having the normal form  $\ulcorner \Phi \text{ denotes } \Psi \urcorner$ . In such sentences, ‘denotes’ is used commissively or in an ontologically committing way, where a sentence of the form  $\ulcorner \Phi \text{ denotes } \Psi \urcorner$  entails  $\ulcorner \Psi \text{ exists } \urcorner$ .<sup>12</sup> Our

---

<sup>11</sup> If you think a true “reductio” ends in a formal contradiction, add the denial of (3) as a further premise. Similarly, in section 5, add the denial of (3\*).

<sup>12</sup> This would suggest that, e.g., ‘Pegasus’ denotes Pegasus’ ontologically commits you to Pegasus. Yet the arguments below do not utilize fiction-terms, so in one respect, the matter will be irrelevant. (The arguments do not concern so much whether the object of a term *exists*, but rather the metaphysical *nature* of an existing object.) It is a bit misleading, then, to say that the arguments concern the “ontology” of a language; more precisely, they concern the *metaphysics* of the objects that our terms denote. (Thanks to an anonymous referee for pressing me to clarify this.) Even so, since it is more customary and succinct to talk of the “ontology” of a language, I shall continue to speak thusly. (For more on the distinction between ontology and metaphysics, see Achille Varzi, ‘Doing Ontology without Metaphysics’, *Philosophical Perspectives*, 25 (2011): 407-423.)

Truth be told, I suspect the new rule following argument also bears on a question like “What, if anything, does ‘Pegasus’ denote?” But to keep my task manageable, I do not defend this here. Yet briefly, the difficulty would be to rule out a non-standard isomorphism where (say) ‘Pegasus’ denotes nothing’ is true iff ‘Pegasus’ denotes nothing *concrete*, yet still denotes something abstract. (In such a model, quantification in negative existentials would be restricted to concreta; even so, each term would denote something abstract or concrete.)

topic thus concerns knowledge of these interpretations, or rather, explicit knowledge of this sort, in contrast to the varieties of tacit or implicit knowledge. What is the explicit/implicit distinction exactly? I cannot say. But for present purposes, we may just stipulate that “explicit” knowledge that  $p$  is knowledge which (ceteris paribus) suffices for an ability to express ‘ $p$ ’ as an answer to a question  $Q$ .<sup>13</sup> Alternately, I call this “knowing an answer to  $Q$ ,” where it is understood that such knowing is explicit knowing. In all cases, however, ‘answer’ is used in a normative sense, so that an “answer” is the same as an adequate answer. At minimum, such will be a true assertion offered as a reply to a question, which provides the kind of information that the inquisitor wanted. (This will become clearer as we go.)

Thus, the idea is to switch the topic of the rule-following argument to explicit knowledge of interpretations of the language, in contrast to a speaker’s ordinary knowledge of the language. In that case, the rule following argument suggests that for the purposes of fundamental ontology, knowing an answer to ‘What does ‘ $\Phi$ ’ denote?’ cannot consist in explicit knowledge of interpretations.<sup>14</sup> However, there is the rub, for knowing such an answer is defined as explicit knowledge of an interpretation. So ultimately, the worry is that it is impossible to know an answer to ‘What does ‘ $\Phi$ ’ denote?’ when asked for such ontological purposes.

---

<sup>13</sup> On “abilities,” see R. Millikan *On Clear and Confused Ideas* (Cambridge: Cambridge University Press, 2004), ch. 6. Strictly speaking, my symbol ‘ $p$ ’ is ill formed, since if ‘ $p$ ’ is replaced with an atomic expression, no concatenation takes place. But for convenience, I assume that whatever expresses  $p$  will be a non-atomic expression.

<sup>14</sup> J. Schaffer “Knowing the Answer,” *Philosophy and Phenomenological Research* 75, no. 2 (2007): 383–403. argues that “knowing the answer” to a question  $Q$  always requires more than just knowing  $p$ , where  $p$  happens to answer  $Q$ . Rather, one must know  $p$  as the answer to  $Q$ . If preferred, this view can be adopted in what follows. The new rule-following problem gets going regardless of whether Schaffer is right. For the problem is that no interpretation counts as an answer in the first place; no interpretation communicates the right information.

Thus the ambition of “metaphysical semantics” is under threat. The goal of such a semantics is to answer how the language relates to fundamental reality. But the problem is that it may be impossible to know the answers to questions of basic ontology. Such a quest would be in vain.

Before moving on, five caveats are in order. First, I ultimately do not endorse the pessimism just described. (Even so, the problem is worth examining, for it brings us to a philosophically notable “skeptical solution.”) Second, I wish to declare neutrality on whether the argument developed here corresponds to anything in Wittgenstein or Kripke. Third, I shall assume (if only for convenience) that one can discuss rule following arguments independently of private language issues. (The argument may well be connected to matters concerning private language, but I do not wish to investigate that here.) Fourth, although the new rule-following problem is eventually resolved, it is notable that the pessimism is still relatively modest: It bears only on questions with the normal form ‘What does ‘ $\Phi$ ’ denote?’’. Consequently, even if such questions are problematic, it would not follow that all metaphysical questions are. (I shall say more about this at the end.) Fifth, and relatedly, the paradoxical conclusion is limited to cases where such questions are raised for specific metaphysical purposes. Hence, it has no bearing when a linguist asks ‘What does ‘ $\Phi$ ’ denote?’’ if the aim is just to determine equivalence relations, model compositionality in the language, or the like. The legitimacy of formal semantics, as a sub-discipline of linguistics, is not at issue.

## *2. Introduction to the Problem*

But prototypically, when the metaphysician asks ‘What does ‘ $\Phi$ ’ denote?’’, it is not for linguistic purposes as such; instead, the aim is to know about the basic ontology “behind” the language.

Still, it is clear that an answer (if any) to ‘What does ‘ $\Phi$ ’ denote?’<sup>1</sup> would have the normal form ‘‘ $\Phi$ ’ denotes  $\Psi$ ’<sup>1</sup>, i.e., it would be an interpretation. But it should also be uncontroversial that some interpretations could not answer any questions; I have in mind homophonic interpretations such as:

- (a) ‘Chair’ denotes chairs.
- (b) ‘Table’ denotes tables.
- (c) ‘Kripke’ denotes Kripke.

After all, without further comment, these are uninformative. The same complaint would hold against the following:

- (d) ‘Chair’ denotes what it denotes.<sup>15</sup>

To be sure, none of these are logically trivial. It is contingent that the sign-design ‘chair’ denotes anything at all. But the fact remains that (a)-(d) are uninformative to English speakers (and unintelligible to non-English speakers). Hence, they are not suited to answer anyone’s questions.

To be clear, homophonic interpretations might carry lots of information in Dretske’s sense of “carrying information.”<sup>16</sup> But they remain uninformative roughly in the familiar, Fregean sense, when ‘Hesperus = Hesperus’ is declared “uninformative.”<sup>17</sup> In the case of ‘Hesperus = Hesperus’ too, the sentence may carry lots of Dretskean information—yet it is still uninformative in an obvious respect, in a way that makes it ill-suited to answer anyone’s questions. It is in roughly the same sense that (a)-(d) are uninformative.

---

<sup>15</sup> Cf. p. 350 of H. Field, “Tarski’s Theory of Truth,” *Journal of Philosophy* 69, no. 13 (1972): 347–375.

<sup>16</sup> See Fred Dretske, *Knowledge and the Flow of Information*, (Cambridge, MA: MIT Press, 1981).

<sup>17</sup> I say “roughly,” since unlike homophonic interpretation, ‘Hesperus = Hesperus’ looks uninformative partly because it is logically necessary. Yet ‘‘Hesperus’ denotes Hesperus’ is “uninformative” in a similar way, even though it is non-necessary. Both sentences remain unsuited to answer anyone’s questions.



Regardless, if information is what's desired, this can seem easy enough to accommodate:

(5) 'Chair' denotes single-seat furniture with a back-rest.

Indeed, a heterophonic interpretation can answer a question about denotation in many contexts. Yet not in all. As a mundane sort of example, suppose a person asks "Which star does 'Pollux' refer to?," where her purpose in asking the question is to locate the star overhead. Then, 'Pollux' refers to Polydeuces' or 'Pollux' refers to the red giant in Gemini' may not answer the question—even if they teach the inquisitor something new. That's because the person may be equally unsure where to pinpoint overhead the denotation of 'Polydeuces' or 'the red giant in Gemini'.

More generally, to answer what a term denotes, it is not enough to offer a true interpretation which is informative in some way or other. It must be informative in the right way. This at least means: It must communicate to the inquisitor the specific information she seeks. Thus in the 'Pollux' example, an interpretation using 'Polydeuces' or 'the red giant in Gemini' does not answer the question since it does not communicate the desired information, viz., where the star is located in the night sky.<sup>18</sup>

The same sort of point holds if, instead of stargazing, the aim is to discern the fundamental ontology of the language. Again, not just any true, informative interpretation will answer what 'chair' denotes for that purpose. I will offer two illustrations of this. The more intuitive example concerns Berkeleyan idealism—yet since few see idealism as a live issue, I

---

<sup>18</sup> This indicates the *purpose-relative* nature of wh-questions, per S. Boër & W. Lycan, *Knowing Who* (Cambridge, MA: MIT Press, 1986). Intuitively, if the inquisitor's purpose is to locate Pollux overhead, *x* answers her question only if it enables her to do this. Yet in the present case, we need not discuss purpose-relativity in detail, and so I mention it here only.

shall be brief, using the example just to convey a basic sense of the problem-space. Dialectical details will come with the second example.

Thus, imagine Bishop Berkeley started his philosophical career by asking “What does ‘chair’ denote?” His purpose in raising the question is to know whether the ontology of the language consists in ideas in the mind of God, or of material, mind-independent objects. In such a case, the interpretation at (5) would not answer his question, even if it is accepted as true. It does not provide the information that Berkeley wishes to know, for (5) remains silent on whether the relevant bits of “furniture” are ideal or material objects.

Even so, suppose we supplemented (5) with:

(6) ‘Furniture’ denotes material objects [of such-and-such type].

If (5) and (6) were granted, would that count as answering Berkeley’s question? It might. But oddly, it also might not. That’s because there is a way of reading (6) so that it is *compatible* with idealism. After all, an idealist apparently holds that *all* terms for putative material objects really denote divine mental objects. And ‘material object’ is itself a term for putative material objects! So it follows from the view that ‘material object’ in (6) also maps onto divine mental objects. (Indeed, this is partly how Berkeley was able to sell idealism as the view of the “folk.”) So given the possibility of an idealist interpretation, even if (5) and (6) are assumed true, the initial question may remain unresolved.

This can suggest, moreover, that the question may remain open indefinitely, no matter what is said in reply. After all, if the question lingers even when the object is explicitly said to be “material,” it is hard to see what could resolve the matter. The question seems unanswerable if any materialist answer is open to an idealist-friendly interpretation. In such a case, ‘Chair’ denotes  $\Psi^1$  answers the question only if one has an answer to whether  $\Psi$  is interpreted in the

idealist-friendly way. And so for Berkeleyan purposes, one apparently must also know the answer to what  $\Psi$  denotes, which requires explicitly knowing a third interpretation, along with its interpretation, and thus a regress. So knowing an answer starts to look impossible.

### *3. Metamereology*

Again, if an interpretation is to answer a question, the interpretation must convey the *right* information, i.e., the information desired by the inquisitor. The problem in the Berkeleyan case is that the right information seems incommunicable, if any response is open to both the materialist and idealist interpretations. Again, I will eventually present a reply to the problem, but it is first worth appreciating that the problem is not unique to idealism. Whenever the fundamental ontology of the language is in question, an answer can seem forever out of reach.

As a more contemporary illustration, consider van Inwagen's mereological nihilism.<sup>19</sup> (I do not mean to single out Professor van Inwagen; his nihilism just affords a particularly clean illustration of the broader point.) Qua nihilist, van Inwagen asserts that "chairs do not exist," when 'chair' is used as a term for a composite object, i.e., something "over and above" the parts. Yet he also regards 'Chairs exist' as true in ordinary English. That's because 'chair' in ordinary English is interpreted vis-à-vis the nihilist ontology, where it just denotes "simples arranged chair-wise." Accordingly, 'chairs exist' comes out true in ordinary English, given that the chair-wise arrangements indeed exist.

However, once the different possible interpretations are raised, it could be difficult to settle 'What does 'chair' denote?' for the purposes of fundamental ontology. To bring this out, let me borrow the character of Socrates, with apologies to Plato (and to Socrates for that matter).

---

<sup>19</sup> *Material Beings* (Ithaca, NY: Cornell University Press, 1990).

Suppose Socrates asks us ‘What does ‘chair’ denote?’ to learn specifically whether van Inwagen’s nihilism is correct. Assume, moreover, that Socrates flatters us in trusting our judgment completely; he will not doubt the truth of what we say. And suppose that there is a wholly determinate fact of whether ‘chair’ denotes a composite object or not. Then, even if (5) is a truth offered in reply, his question may remain unresolved.

Imagine that Socrates has studied his van Inwagen, and thus is aware that ‘furniture’ in (5) could<sup>20</sup> conceivably denote nothing but simples arranged in various ways. Then, even though (5) expresses a determinate fact (and Socrates trusts that it does), his question is not yet resolved. For in context, he can appreciate that the truth of (5) still leaves open the nihilist and the non-nihilist possibilities. Yet if both remain live possibilities for Socrates, (5) does not transmit the desired information, viz., whether ‘chair’ denotes a composite or not.

Naturally, we could supplement (5) with the following to clarify:

(7) ‘Furniture’ denotes a composite [of such-and-such type].

But oddly, even granting the truth of (7), both nihilist and non-nihilist readings can remain live, and for much the same reason. The question concerns terms for putative composites, yet ‘composite’ in (7) is itself a term for putative composites! (Note the parallel with Berkeley on ‘material object’.) So for all Socrates knows, (7) may just map ‘composite’ to a nihilist ontology of simples in various arrangements. Further, being an astute fellow, Socrates is aware of this. So even trusting that (7) along with (5) are true, he is still unclear if ‘chair’ denotes a composite.

---

<sup>20</sup> A modal such as ‘could’ in this context expresses some kind of epistemic or doxastic possibility relative to Socrates, rather than metaphysical possibility. But this is expected; epistemic/doxastic possibilities are the relevant possibilities when it comes to answering questions.

Now you might complain that Socrates is just being uncooperative. He ought to interpret our words in the normal way, unless there is special reason to do otherwise. But Socrates is unsure which way is the “normal” way. Granted, if the ontology of the language is non-nihilistic, then ‘composite’ in (7) normally denotes a *composite*. Yet when asking his question, Socrates is the perfect agnostic. More, he realizes that if van Inwagen is right, terms for composites are “normally” mapped onto a nihilistic reality. Hence since Socrates does not prejudge the nihilism issue, he is equally agnostic on which interpretations count as the “normal” ones.

There is more to say here, but at least one can now feel the threat of regress. If (5) and (7) are insufficient, we might add:

(8) ‘Composite’ denotes a whole that is not the sum of its parts.

Yet here too, the problem is that the entire clause to the right of ‘denotes’ might have either interpretation. For this clause is also a term for putative composites. But if so, van Inwagen’s view already implies that its “normal” interpretation maps it to mere simples in various arrangements. And since Socrates’ question does not prejudge these matters, both the nihilistic and non-nihilistic interpretations will be live options. So again, (8) will not resolve his question.

Granted, if the semantics is compositional, the nihilist reading of the clause requires its terms to be semantically “aligned” in certain ways, if the truth of (8) is a constant. But such is entirely possible. (E.g., ‘whole’ and ‘sum of its parts’ could switch denotations, so the clause means “any plurality of simples that is not a composite.”<sup>21</sup>) Indeed, non-standard isomorphisms

---

<sup>21</sup> In a nihilistic domain, the reinterpretation means ‘the sum of its parts’ will be referentially empty. But this creates no problem; (8) will still have ‘composite’ denote pluralities of simples. A given plurality will satisfy ‘whole’ under the reinterpretation, yet the plurality will not be identical to any  $x$  that is “the sum of its parts”—simply because no  $x$  satisfies ‘the sum of its parts’ under the reinterpretation.

are always possible, as Putnam is fond of noting.<sup>22</sup> But here, Putnam's model-theoretic considerations are not supporting an ontological point like referential indeterminacy or metaphysical anti-realism. Rather, they buttress a certain metaontological point, viz., that Socrates' ontologically-driven question may remain unanswered, even if the semantics is compositional and our responses are true. (The point concerns ontological inquiry, rather than ontology per se.)

Observe that Socrates will also be unsatisfied by interpretations that deploy a demonstrative, as in:

(9) 'Chair' denotes *that*.

If (9) is accompanied by a suitable ostension, it seems to provide a "nonconceptual information link" that "directly acquaints" Socrates with the object, in a way that dispels the question.<sup>23</sup> Yet if it is genuine question whether composite objects are "out there" to denote, (9) merely pushes the issue back onto the demonstrative.

The problem is not that the meaning of (5)-(9) is indeterminate. It is assumed that they are determinately true, that their terms determinately refer, and that they determinately contradict one of nihilism or non-nihilism. One could even invoke "reference magnetism" to explain all this. For the problem is not semantic, but more epistemic. It is that our responses fail to convey the right information to one who lacks it. The problem is like the stargazer's: Even though it is perfectly true that 'Pollux' denotes the red giant in Gemini, this may not direct the inquisitor where to look. Analogously, even though the interpretations may express truths that contradict

---

<sup>22</sup> See H. Putnam, "Realism and Reason," *Proceedings and Addressess of the American Philosophical Association* 50 (1977): 483–498; "A Problem About Reference," in *Reason, Truth, and History* (Cambridge: Cambridge University Press, 1981), 22–48; "Models and Reality," *Journal of Symbolic Logic* 45 (1980): 464–482.

<sup>23</sup> Cf. ch. 6 of G. Evans, *The Varieties of Reference* (Oxford: Clarendon Press, 1982).

(say) nihilism, they may not communicate as much, given Socrates' assumptions (or rather, lack thereof). In both cases, a gap in the inquisitor's background thwarts a serviceable understanding of our reply. And in Socrates' case, it seems that *any* reply to the question will fail as an answer.

Speaking broadly, Socrates' question about 'chair' occurs in a context where the basic ontology of the entire language is in question. For that reason, any 'chair'-interpretation will use a term whose denotation is equally in doubt; the nihilism issue will pertain to the right-hand term just as much as 'chair'. So even granting their truth, the interpretations are "just more theory." Or in the Wittgensteinian idiom, the interpretations will "hang in the air" along with what they interpret. To settle Socrates' question, it appears the interpretations themselves need to be interpreted, and thus a regress. In which case, there is no answering Socrates' question.

We certainly have explicit knowledge of interpretations; that is not in question. But we might have thought this constitutes knowledge of answers to questions about fundamental ontology. Yet the rule following problem suggests that, whatever this knowledge amounts to, it does not amount to knowing the answer to Socrates' type of question.

#### *4. Objections and Replies*

Before offering my own response to the problem, I wish to consider two others that end up inadequate. The first is that the interpretation at (7) should not be seen as a sentence of ordinary English. Rather, it should be seen as a sentence of "Ontologese," to borrow Sider's term.<sup>24</sup> For if

---

<sup>24</sup> T. Sider, "Ontological Realism," in *Metametaphysics*, ed. David Chalmers, David Manley, and Ryan Wasserman (Oxford: Oxford University Press, 2009), 384–424; *Writing the Book of the World*, op. cit; "Hirsch's Attack on Ontologese," *Noûs* (forthcoming), [http://tedsider.org/papers/hirsch\\_attacks.pdf](http://tedsider.org/papers/hirsch_attacks.pdf).

Socrates' question pertains only to ordinary English, the Ontologese sentences would not be at issue—in which case (7) qua Ontologese can indeed answer his question.

However, Socrates cannot help it if he has questions about Ontologese; it is *ex hypothesi* not his native tongue. Yet given Ontologese-to-English translations, Ontologese is then indirectly indicted by the question. For the translations would be interpretations whose right-hand term is a term of English, and the basic ontology of English is what's at issue.

The second reply is that the regress problem is self-undermining. For the problem arises by first articulating the nihilist and non-nihilist interpretations—and in that process, 'composite' is unequivocally understood as a term for composites. Thus, insofar as Socrates grasps his own question, his understanding of 'composite' must already be unproblematic. So it seems the regress should stop with (6).

The reply is roughly that, by his own lights, Socrates has already said what supposedly cannot be said. But in response, one might take Lewis as inspiration when he writes:

I could have said my say fair and square, bending no rules. It would have been tiresome, but it could have been done. The secret would have been to resort to 'semantic ascent' ... [But] If you want to hear my story told that way, you probably know enough to do the job for yourself.<sup>25</sup>

Indeed, it is familiar that self-refutation is often avoidable by semantic ascent. The object-language/meta-language distinction may strike some as contrived, though it can drawn in a principled way as Kripkenstein does, between past utterances vs. present ones. Or, one could suppose that Socrates' question occurs in thought. Then, the various interpretations could be thought by Socrates, even if these are things that cannot be said.

---

<sup>25</sup> P. 41 of D. Lewis, "Elusive Knowledge," *Australasian Journal of Philosophy* 74, no. 4 (1996): 37–41.



The distinction would be unlike the distinction between English and Ontologese, since here Socrates can assume he understands the expressions of both “languages.”<sup>26</sup> Still, the distinction would allow him to raise a non-defective question about an English expression, while taking the “language” of the question for granted. Nevertheless, in entertaining this question in thought, no sentence of English could act as an answer. For the ontology of any English interpretation would be in question as much as the term it interprets. So no English interpretation could resolve what an English term denotes, for Socrates’ ontological purposes.

### 5. *A Skeptical Solution*

Nevertheless, the second reply points the way to a better one. It seems that there can be an answer to Socrates question—namely, when the answer occurs in the very language of the question. So for instance, if Socrates’ question occurs in thought, he could answer it with his subsequent thought that ‘chair’ denotes a composite. For his concept of a composite must be understood as representing composite objects, if Socrates is to grasp his own question. Yet then, this same concept can function as an answer.

We might reconstruct all this in a way that parallels the Wittgensteinian argument (where a starred statement is the analogue of the unstarred correlates):

- (1\*) Knowing an answer to ‘What does ‘ $\Phi$ ’ denote?’<sup>1</sup> when asked for Socrates’ ontological purposes (i.e., knowing an answer to “ $Q_\Phi$ ”) requires explicitly knowing a true sentence with the normal form ‘ $\Phi$ ’ denotes  $\Psi$ ’<sup>1</sup> that communicates the right information. [Assume]

---

<sup>26</sup> Henceforth, for convenience, I use ‘language’ and related expressions like ‘term’ in relation to any system of linguistic *or mental* representations.

(2\*) 「‘ $\Phi$ ’ denotes  $\Psi$ 」 communicates the right information only if Socrates knows an answer to 「What does ‘ $\Psi$ ’ denote?」 when asked for his ontological purposes (“ $Q_\Psi$ ”).

[Assume for reductio]

(2.1\*) Then, knowing an answer to  $Q_\Phi$  requires knowing an answer to  $Q_\Psi$ .

[From (1\*), (2\*)]

(2.2\*) And knowing an answer to  $Q_\Psi$  requires knowing an answer to  $Q_\Omega$  for a term  $\Omega$ .

[From (1\*), (2\*)]

(2.3\*) And knowing an answer to  $Q_\Omega$  requires knowing an answer to  $Q_\Delta$  for a term  $\Delta$ .

[From (1\*), (2\*)]

(Etc...)

(3\*) So, knowing an answer to  $Q_\Phi$  requires explicitly knowing infinitely many

interpretations.

[From the preceding]

(4\*) So, there is a way of knowing an answer to  $Q_\Phi$  that does not require knowing an

answer to  $Q_\Psi$ .

[By reductio, (2\*)-(3\*)]

As in Wittgenstein’s argument, the regress is blocked by denying the second premise. Though we saw that Socrates may understandably want interpretations of his interpretations. Again, if (5)-(9) rest on an uninterpreted term, then for all he has been told, the language might still map onto a nihilist ontology, an idealist ontology, a Pythagorean ontology, or what have you.

Even so, the suggestion is that an interpretive-ontological question can be answered if the reply occurs in the very language of the question. ‘Composite’ within the question must be understood as denoting composites, if the question is to be grasped in the first place. So the term can be understood unproblematically if offered as an answer.

Regardless, there is a concession to be made. Even when an interpretation acts as an answer, a deeper ontological question will go unaddressed. For instance, if Socrates answers his question in thought, a question about the ontology of thought is yet to be raised. That question seems like the more important one, moreover, since its answer would supersede the first answer. It would supersede it in the sense that its answer could inform him whether the first answer was really a non-nihilistic answer. (If thought itself is mapped to a nihilistic ontology, then what seemed like a non-nihilist judgment about an ontology of “composites” might end up being a nihilistic one.)

That the further questions have superseding answers is important. For if one cares to interpret ‘chair’ in the first place, then there is some interest in interpreting its interpretation. After all, the object of ‘chair’ depends not only on how the term is interpreted, but also on its interpretation is interpreted (and on how the interpretation of its interpretation is interpreted, and so forth).

#### *6. Metaphysical Semantics Revisited*

The prospects for Sider’s “metaphysical semantics” thus seem limited. Some terms will inevitably go uninterpreted (or the interpretations will circle on themselves, which also bodes ill for answering questions). So it seems there will always be superseding ontological questions which go unanswered—even if we are (for whatever reason) appeased at some stage of the regress. Of course, the problem is not unique to Sider. Insofar as any answer to an interpretive-ontological question is possible, it will be an answer that ignores other questions with superseding answers.

Sider himself raises an “ideological” as well as a “propositional” regress to challenge his metaphysical semantics.<sup>27</sup> Yet these are different from the rule following problem. The ideological regress is driven by the desire for a more fundamental vocabulary in which to describe reality. Sider replies here that there is an “ideological ground floor:” It is just a fact that some atomic terms of the language, e.g., ‘structure’ and perhaps ‘electron’ or ‘neutrino’, are the most fundamental terms we have coined thus far. (It may be unclear what qualifies a term as “the most fundamental,” but Sider’s reply seems solid.)

Secondly, Sider’s propositional regress arises by there always being more fundamental facts to describe. Yet Sider notes it is possible to describe an infinite set of facts, of ever increasing fundamentality, in just finitely many terms. E.g., if the world turns out to be mereologically “gunky,” one can describe the infinite division of parts recursively. If we start with an object weighing one gram, our theory can say it decomposes it into two parts weighing 0.5 grams apiece, which in turn decompose into parts weighing 0.25 grams, etc. More broadly:

(\*) If there is an object weighing  $m$  grams, then it is a composite of 2 elements, each weighing  $m/2$  grams.

So it is not as if one would need to utter infinitely many sentences in order to describe an infinite series of facts.

Even so, the new rule following problem is not driven by the desire for more fundamental terms, nor for an infinite number of facts. Rather, it owes to any interpretation being interpretable with respect to a variety of ontologies. Thus, suppose we settle on a set of terms as our most

---

<sup>27</sup> 2011, p. 133-134. Sider also notes a third, mereological regress. But as far as I can tell, it is a special case of the propositional regress, re: propositions about parts and wholes. (Indeed, Sider’s example of the propositional regress is mereological; see p. 134.)

fundamental language, ('structure', 'electron', etc.), so that the ideological regress is not a factor. And suppose our theory recursively describes the world as breaking down into more and more fundamental particles, in the manner of (\*). Then, there is still a Socratic question in the area. Since (\*) contains the word 'composite', Socrates might ask the same kind of question that arose with (6): Does 'composite' in (\*) denote *composites*? Or does it just denote "heaps" of simples?<sup>28</sup> (Similarly: Does 'electron' denote a composite?) So even once the ideological and propositional regresses have been halted, the new rule following problem remains.

Sider might retort that to interpret 'electron' for ontological purposes is *ipso facto* to continue the ideological regress. But if so, then I would withdraw the earlier agreement that there is an "ideological ground floor." Questions about the language do not end with the primitive vocabulary. The interpretive-ontological questions equally exist for primitive terms, though their "primitive" status means such questions are to remain unanswered.<sup>29</sup>

## 7. *Quietism*

If an interpretation ends the rule-following regress, it is not because it answers the most basic ontological questions. Rather, it is because the interpretation is used "blindly," as it were. The

---

<sup>28</sup> One might object that (\*) patently describes the world as non-nihilistic, given that it describes it as gunky. But in fact, if the interpretation of 'composite' is in question, then it is also in question whether it is truly describing *gunk*. It could instead describe a world where no composition occurs, and where the mereological atoms are infinitesimals. In that case, (\*) separates out heaps into smaller and smaller piles, ad infinitum.

<sup>29</sup> Perhaps Sider's regimented quantifier, expressed by '**existence**' in his (2009), can help resolve Socrates question? It might be defined as unequivocally denoting what is actual, where this determines an extension of (say) non-composites. But unfortunately, there are difficulties with '**existence**' as well; see T. Parent, "Ontic Terms and Metaontology—or, On What There Actually Is," *Philosophical Studies*, 170 (2014): 199-214.

right-hand term will be one that we understand, but this understanding will occur absent an answer to what it denotes. Or in Quine's vocabulary, it will be part of a "background language"—a language where semantic questions are answered by terms which are themselves used unquestioningly.<sup>30</sup> (But n.b., Quine's regress is importantly different; it arises from his behaviorism and semantic nihilism.) Nonetheless, even if the background language affords answers, they function as such only if we ignore the superseding question about their interpretation. (Quine calls this "acquiescing in our mother tongue" where our terms are "taken at face value." One might also say it is acquiescing in the linguistic "practices" or "customs.")

This can encourage deflationism about ontology since, as Jay Rosenberg once said, "[t]o terminate a regress of background languages in our own 'taken at face value' is to cash a forged check with counterfeit currency"<sup>31</sup> Yet if Socrates' question can result only in "counterfeit currency," this does not mean that ontology is shot. The deflationism would be limited to ontological questions phrased in semantic terms. And this leaves untouched non-semantic questions of the form 'What is  $\Phi$ ?'. For all that has been said, the physical sciences answer these questions just fine. One may protest that the deflationary point extends to those answers as well, since they ultimately occur in a language where interpretive questions go unanswered. Perhaps—but the regress suggests is a mistake to pursue such questions, as if one could answer more fundamental ontological questions. (Indeed, a regress also results if Socrates' inquiry concerns whether deflationism or realism is correct.)<sup>32</sup>

---

<sup>30</sup> "Ontological Relativity," *Journal of Philosophy* 65, no. 7 (1968): 185–212.

<sup>31</sup> P. 56 of his *Linguistic Representation* (Kluwer: Dordrecht, 1974).

<sup>32</sup> In contrast, 'What is an electron?' need not motivate a regress, since normally it does not put in question *any* term co-referring with 'electron'. Normally, it is a request for a theory of electrons, one that provides informative, co-referring descriptors.

So the lesson here is not so much deflationism as *quietism* about interpretive questions, when asked for ontological purposes. Yet quietism does not mean that an analysis of language has nothing to offer. The philosopher can still help unravel the conceptual knots that arise in our theories about the world, e.g., in talk of “holes.” She can assist by paraphrasing such talk into clearer language, and getting clearer on things is genuine progress. Though again, the paraphrases should not be seen as answering the most fundamental ontological questions. Philosophy can at least clarify what the sciences say, and in this, perhaps philosophy of science is philosophy enough. But what the regress signals that the science per se is ontology enough—or rather, that it is about all the ontology one should expect.<sup>33</sup>

#### *Appendix: Compare and Contrast*

It is appropriate to address other parts of the metaontology literature, though it is easiest to do so after my own position is in view. Hence, the appendix.

First, the new rule following argument bears some similarity to the Langton-Lewis argument for “Kantian humility.”<sup>34</sup> Still, there are clear differences. Most basically, the Langton-Lewis argument concerns knowing intrinsic natural properties only, and the problem is

---

<sup>33</sup> There is a kinship here with Hofweber’s stance against “immodest metaphysics.” We agree it is futile to try to discern the ontology “underlying” the language of science, as if metaphysicians could provide answers that supersede those of science. See Thomas Hofweber, “Ambitious yet Modest Metaphysics,” in *Metametaphysics*, D. Chalmers, D. Manley, & R. Wasserman (eds.), (Oxford: Clarendon Press, 2009), 260-289.

<sup>34</sup> See Langton, R. *Kantian Humility*, (Oxford: Oxford University Press, 1998); “Elusive Knowledge of Things in Themselves,” *Australasian Journal of Philosophy* 82.1 (2004): 129-136; also, D. Lewis, “Ramseyan Humility,” in *Conceptual Analysis and Philosophical Naturalism*, R. Nola & D. Braddon-Mitchell (eds.), (Cambridge, MA: MIT Press, 2001/2009), 203-222.

motivated by means of causal powers. (Their point is that we are not acquainted with natural properties, but only their causal effects.) However, the new rule-following problem assumes the existence of neither intrinsic natural properties nor causal effects—much less is one used to problematize the other. Accordingly, the scope of the rule following problem is wider. It arises equally for ‘electron’, ‘neutrino’, or whatever terms are thought to pick out the perfectly natural properties. It even arises for the term ‘causal effects’ as applied to perception. Socrates might equally ask for his ontological purposes what ‘perceptual causal effects’ denotes. Sense data? Disjunctivist contents? Again, the question seems answerable only relative to a background language. (Also, Lewis at least draws no quietist conclusions.)

What’s more, I am not aligning with Hirsch despite all my talk of interpretation.<sup>35</sup> For one, the present discussion is centered on alternative interpretations of terms rather than of quantifiers. For another, I make no demand that interpretive charity requires us to regard ontologists as talking past each other. (And what sort of charity is that?<sup>36</sup>) Third, there is no “superficialist” recommendation, where ontological questions are to be answered according to the ordinary usage of English quantifiers. Nevertheless, like Hirsch, I appeal to alternative interpretations to “deflate” the significance of ontological debate. But again, my sort of deflation is not that ontological disputes are merely verbal disputes. Rather, the “deflation” occurs because

---

<sup>35</sup> E. Hirsch, “Quantifier Variance and Realism,” *Philosophical Issues* 12, no. 1 (2002): 51–73; “Physical-Object Ontology, Verbal Disputes, and Common Sense,” *Philosophy and Phenomenological Research* 70, no. 1 (2005): 67–97; “Ontological Arguments: Interpretive Charity and Quantifier Variance,” ed. T. Sider, J. Hawthorne, and D. Zimmerman (Oxford: Wiley-Blackwell, 2007), 367–381; “Ontology and Alternative Languages,” in *Metametaphysics*, ed. D. Chalmers, D. Manley, and R. Wasserman (Oxford: Clarendon Press, 2009), 231–259.

<sup>36</sup> I refer here to the criticism of Hirsch on pp. 59–60 of Hawthorne, “Plentitude, Convention, and Ontology,” in his *Metaphysical Essays* (Oxford: Oxford University Press, 2006), 53–69.



the superseding questions inevitably go unanswered. And the consequence of this is not superficialism but rather a kind of resignation or “quietism.”

Let me also repeat I am not arguing for the inscrutability of reference or metaphysical anti-realism, despite mention of Putnam’s model-theoretic argument. Recently, there have been some impressive attempts to resuscitate Putnam’s argument from Lewis’ “naturalness” objection.<sup>37</sup> Regardless, the present piece is not furthering that effort. It is instead assumed throughout that ‘chair’ has its “normal” or “standard” interpretation, and that there is a (theory-independent) fact of the matter about what that is. The threat was not anti-realism, but rather a kind of ignorance. It was an epistemic problem about answering the question of what ‘chair’ denotes for ontological purposes. In Socrates case, it seemed that nothing could settle the question, since as far as he was concerned, any reply was open to more than one interpretation. So the relevant knowledge remained missing, even if the facts are wholly determinate.

But now it is best to separate this from Bennett’s skeptical metaontology.<sup>38</sup> Bennett thinks there are epistemic limits on answering ontological questions as well, yet her skeptical stance has more to do with insufficient evidence. In contrast, the present view is not highlighting an evidential deficiency. One can suppose that (5)-(9) are not only true, but also that their

---

<sup>37</sup> D. Lewis “Putnam’s Paradox,” *Australasian Journal of Philosophy* 62, no. 3 (September 1984): 221–236. The rejoinders I have in mind are: R. Williams, “Eligibility and Inscrutability,” *Philosophical Review* 116, no. 3 (2007): 361–399; L.A. Paul, “Realism About Structure and Kinds,” in *The Metaphysics of Science*, ed. Stephen Mumford and Matthew Tugby (New York: Oxford University Press, 2012), 203–221; Tim Button, *The Limits of Realism*, (Oxford: Oxford University Press, 2013).

<sup>38</sup> K. Bennett, “Composition, Co-location, and Metaontology,” in *Metametaphysics*, D. Chalmers, D. Manley, and R. Wasserman (eds.) (Oxford: Clarendon Press, 2009), 38–76.

evidential probability is as high as you like. The problem, however, is that there are always superseding ontological questions which are not even articulated, much less answered.

McDowell and Wright also connect quietism to rule following considerations.<sup>39</sup> Yet it is Kripke's version of the argument which is relevant there, and the issue is whether key terms are indeterminate in meaning.<sup>40</sup> But again, semantic indeterminacy is not the topic here. It consists rather in epistemic worries of a sort, about using the language to talk about that very language.<sup>41</sup>

One might think of the different views here in terms of different "gaps." An anti-realist holds that there is a gap in the facts (incl. facts about reference). The skeptic, on the other hand, stresses a gap in the evidence. Whereas Kripkenstein worries concern semantic gaps. The new rule following problem, however, emphasizes a certain *informational* gap—or a gap in the set of answers to questions. It brings out that no interpretation conveys the right information to settle what ontology underlies the entire vocabulary. The only answers in the area are relativized to a "background language," where questions about that language's own ontology are ignored.<sup>42</sup>

---

<sup>39</sup> J. McDowell, "Non-cognitivism and Rule-Following," in C. Leich & S. Holtzman (eds.), *Wittgenstein: To Follow a Rule* (London: Routledge, 1981) pp 141-162; "Wittgenstein on Following a Rule," *Synthese* 58, 325-363; also C. Wright, *Truth and Objectivity*. (Cambridge, MA: Harvard University Press, 1992).

<sup>40</sup> This I take to be the key issue in Thomasson also, though she too does not invoke rule following considerations. See A. Thomasson, "Existence Questions," *Philosophical Studies*, 141 (2008): 63-78; A. Thomasson, "Answerable and Unanswerable Questions," in Chalmers et al. (eds.), *Metametaphysics*, 2008, pp. 444–471.

<sup>41</sup> Huw Price is also associated with "quietism," yet he adopts it only in a limited way. See pp. 250-1 of *Naturalism without Mirrors*, (Oxford: Oxford UP, 2011). Also unlike Price, no anti-representationalist stance is here taken.

<sup>42</sup> This paper stems from chapter one of my dissertation (though it is markedly different). I owe thanks to my dissertation director, William G. Lycan, and to my committee: Dorit Bar-On, Thomas Hofweber, Ram Neta, and Keith Simmons. I also thank Derek Haderlie, as well as an audience at Virginia Tech, and an audience at the 4th annual Regional Wittgenstein conference.