

# The Metaethical Insignificance of Moral Twin Earth<sup>1</sup>

J.L. Dowell

Draft at 6/16/14

10,829 words

What methodology is appropriate for semantic theorizing? In particular, what types of consideration place genuine constraints on an adequate semantics for normative and evaluative expressions?<sup>2</sup> Data points widely recognized among linguists include facts about ordinary uses of expressions and competent speakers' judgments about which such uses are acceptable and appropriate.<sup>3</sup> The contemporary literature in metaethics reflects the widespread assumption that there is an additional source of data linguists don't rely upon, however, namely, competent speakers' judgments about the possibility of disagreement with *rival* and *hypothetical* speech communities. Both "rival" and "hypothetical" are both important here. Since the communities in question are both rival and hypothetical, the experiments are ones that, by stipulation, feature a language respondents do not competently speak. The judgments, then, are about whether it is possible for us, using English, to express disagreement with speakers of a different language we are not competent with, when they use sentences in that distinct language.

Thought experiments generating such judgments include Hare's missionary, Smith's "rival speech community", and, most famously, Horgan and Timmons' moral twin earth. While the precise targets of these thought experiments vary, each is thought by

---

<sup>1</sup> Many thanks to audiences at the Arche Centre at the University of St. Andrews, the University of Bogazici, Cornell University, the University of Massachusetts-Amherst, the University of North Carolina-Chapel Hill, Syracuse University and the University of Umea for discussion and especially to Ruth Chang, Jamie Dreier, Tristram McPherson, David Plunkett, Gideon Rosen, Geoff Sayre-McCord, Ted Sider, David Sobel, Will Starr, Nick Sturgeon, and Tim Sundell.

<sup>2</sup> As Michael Smith (1994) notes, one of the central tasks of metaethics is to understand ordinary moral discourse, not terms of philosophical invention.

<sup>3</sup> See also Yalcin (manuscript).

proponents to in some way or other put pressure on purely descriptivist semantics for our moral terms. (By a “purely descriptivist semantics for our moral terms” or “Descriptivism”, I’ll mean any view on which our moral terms receive a semantics of the same type as that of ordinary, descriptive terms, e.g. natural kind terms or empirical terms more broadly.) Typical discussions involve four steps. The first offers a characterization of the use of certain terms in the language of a hypothetical speech community, terms that are, by stipulation, to function in ways similar to our moral terms in English. The second makes a claim about the pretheoretical judgments of ordinary speakers about the possibility of our expressing moral disagreement with that community or about whether our terms are intertranslatable with theirs, while the third step defends the claim that those judgments are incompatible with Descriptivism. The final step draws a conclusion about the semantics of moral terms in English, typically, that this incompatibility means that a Descriptivist semantics for moral terms must be incorrect.

The metaethical literature reflects a variety of replies to these objections, especially to the moral twin earth thought experiments. What these replies share, though, is the assumption that our intuitions in thought experiments of this kind, when formulated properly, have probative value. In other words, existing replies focus on the first three steps; either they argue that the thought experiments have not been properly characterized<sup>4</sup> or that ordinary speakers fail to have the pretheoretical judgments their

---

<sup>4</sup> Merli (2002), Laurence, Margolis, and Dawson (1999). For a reply to the latter, see Rubin (2008).

proponents have claimed<sup>5</sup> or that descriptivism or some recognizable successor is compatible with our intuitions.<sup>6</sup>

Here I focus instead on blocking the anti-Descriptivist conclusion by providing principled grounds for rejecting this shared assumption. Even when formulated properly, I'll argue, we should accord our judgments in these cases no probative value for semantic theorizing.<sup>7</sup>

If correct, this conclusion has important consequences not only for our assessment of the merits of Descriptivism. Many rivals to pure, descriptivist theories count among their advantages the ability to accommodate our judgments in these thought experiments. If we should accord those judgments no probative value, such theories lose an important source of their support.<sup>8</sup>

The plan of action is as follows: First, I'll focus simply on clearly representing the reasoning behind two such thought experiments, Smith's rival speech community and the Horgan and Timmons moral twin earth experiment. Getting clear on this reasoning will help show how the probative value of our pretheoretical judgments about them requires substantive assumptions about what semantic competence with a term consists in, assumptions that are rarely made explicit, let alone defended, by their proponents.

Characterizing these assumptions precisely is the task of section 2.

---

<sup>5</sup> Merli (2002), Sayre-McCord (1997).

<sup>6</sup> Copp (2000), Sayre-McCord (1997), van Roojen (2005).

<sup>7</sup> For a quite different argument for a similar conclusion, see Plunkett and Sundell (2013).

<sup>8</sup> Those who treat accommodating these judgments as a source of support for their positive views include Bjornsson and McPherson (forthcoming), Hare (1952), Ridge (forthcoming), van Roojen (2005), and Wedgwood (2001) and (2007).

Section 3 focuses on how we should understand the notion of semantic competence appropriate here by getting clear on the evidential connection between competent uses of terms and semantic theorizing. Section 4 aims to show just how deeply implausible the required account semantic competence is, given which notion of semantic competence is appropriate for semantic theorizing. Here the argument will have a negative and a positive component. On the negative side, I'll show that the needed account of competence is much too demanding to plausibly fit with what's known about language learning. On the positive side, I will sketch a much less demanding, Descriptivist-friendly account inspired by Ruth Millikan's work, which both fits with the empirical data and is independently well-motivated. That account, as we'll see, accords our judgments about these thought experiments no probative value. While these arguments together don't conclusively rule out all possibility that further information will vindicate their evidential value, they do make this possibility highly unlikely. Given this, we should currently accord those judgments no probative value for semantic theorizing.

### **1. Thought Experiments about Rival, Hypothetical Speech Communities**

Both our two thought experiments target Descriptivist semantics for moral terms, taking casual regulation theories<sup>9</sup> as their sample of such views.<sup>10</sup> Here's the Horgan and Timmons formulation of what's central to such theories:

*CSN Causal semantic naturalism:* Each moral term  $t$  rigidly designates the natural property  $N$  that uniquely casually regulates the use of  $t$  by human beings." (Horgan and Timmons, "New Wave Moral Realism Meets Moral Twin Earth" (reprint of

---

<sup>9</sup> For examples of causal regulation theories, see Boyd (1988), Brink (1989).

<sup>10</sup> To be fair to Smith, it is not clear that he intends for his thought experiment to target any Descriptivist theory other than CSN. He does, though, treat accommodating our judgment about his rival speech community as placing a constraint on a semantics for moral terms and this is a thesis I'll be arguing against. (Thanks to David Plunkett for discussion here.)

1991), 123.)

First, consider Smith's 'rather different community' (1994: 32-3).

“...let's imagine that we come across a rather different community from ours in which they too use **the word 'right'** in a **practice much like moral practice** and let's ask the [the Descriptivist] what we should say about the contents of the **judgments about the rightness and wrongness** of actions made by those in this other community. The problem...is that we cannot *simply assume that the same property is causally responsible for their uses of the word 'right'*. And if it is not, then we have not been given an account of the central platitudes governing our use of that word: namely, the platitude that when A says “x is right” and B says, “x is not right”, then A and B disagree with each other. We have not been given an account of what the word ‘right’ refers to that is consistent with the objectivity of morality.” (The bold emphasis is mine.)

As stated, this argument appears to beg the question against the Descriptivist. Seeing how that's so, however, will help us see how to reformulate it in a way that more clearly captures Smith's thought. To see this, notice that the description of the case builds in the assumption that this 'rather different community' also uses *the word "right"*. It's relatively standard among linguists to individuate words by their meanings, e.g. so that the “bank” that refers to financial institutions and the “bank” that refers to the land at the side of rivers, are distinct words.<sup>11</sup> Given this, on the most natural way of understanding what's meant by “the word ‘right’”, that word is just the English word “right” with its univocal semantics. But in that case, the assumption that they and we both use “the word ‘right’” settles by stipulation that *the word they and we both use* has a single meaning—the univocal English meaning. The trouble is that this assumption, together with the argument's second main assumption, begs the question against the Descriptivist. Its second assumption is that settling sameness of word and so sameness of meaning doesn't settle that what causally regulates their usage is the same as that which causally regulates ours. If that's so, though, there's no need to consult our intuitions about disagreement, as

---

<sup>11</sup> For a distinct way of thinking of word individuation, see Plunkett and Sundell (2013). Thanks to David Plunkett and Tim Sundell for discussion of this issue.

these assumptions alone entail the falsity of CSN.<sup>12</sup> This way of casting the objection leaves it open to a straightforward reply by the proponent of CSN: As soon as you stipulate their word “right” is ours, you stipulate that their term is regulated by the same properties ours is.

While this reply is correct, it’s easy to see how to provide a formulation of Smith’s central thought that keeps clear of this appearance. Suppose we alter his characterization of the case in the following way:

...let’s imagine that we come across a rather different community from ours in which they...use **a word ‘R’** in a practice much like our moral practice; indeed, their term “R” seems to function in their practice in much the way “right” figures in ours. Let’s ask the Descriptivist what we should say about the contents of the **judgments** about actions **expressed using “R”** made by those in this other community. The problem...is that we cannot *simply assume that the property causally responsible for their uses of word ‘R’ is the same as those responsible for our uses of “right”*. And **if it is not**, then we have not been given an account of the central platitudes governing our use of that word: namely, the platitude that when A says “x is right” and B says, “x is not R”, then A and B disagree with each other. We have not been given an account of what the word ‘right’ refers to that is consistent with the objectivity of morality.

This alteration allows us to reconstruct Smith’s argument thus:

1. When we focus on the similarity between our usage of “right” and their use of “R”, we have the sense that, if, on those occasions when we would say of some act that it is right, they would say that it is not-R, they disagree with us. (That is, we have the intuition that the similarities in usage support disagreement.) Call this sense ‘Judgment about Disagreement’.
2. A semantic theory should vindicate Judgment about Disagreement. (Call this claim “Vindicate”.)
3. Descriptivist theories, like CSN, cannot vindicate Judgment about Disagreement (since, according to such theories, we mean different things by “right” and “R”, on the assumption their usage tracks different properties).
4. So, Descriptivist theories, like CSN, are false.

---

<sup>12</sup> Hare’s missionary case suffers from a similar flaw.

Since Vindicate is what's at issue here, I grant #1 for the sake of argument. Before getting to what might be said for or against Vindicate, though, let's put our second thought experiment on the table.

Here's a standard statement of the Moral Twin Earth thought experiment from Horgan and Timmons:

“Now consider Moral Twin Earth, which, as you might expect, is just about like good old Earth: same geography and natural surroundings...people who live in the twin United States by and large **speak Twin English**...Of particular importance here is the fact Moral Twin Earthlings have a vocabulary that works much like human moral vocabulary; **they use the terms “good” and “bad,” “right” and “wrong”** to evaluate actions, persons, institutions and so forth...Let us suppose that investigation into Twin English moral discourse and associated practice reveals that their uses of twin moral terms are causally regulated by certain natural properties distinct from those that...regulate English moral discourse.” (Horgan and Timmons 1993:128-129. Bold emphases mine.)

“Given all these assumptions and stipulations about Earth and Moral Twin Earth, what is the appropriate way to describe the differences between moral and twin moral uses of “good” and “right”? Two...options are available. On the one hand, we could say that the differences are analogous to those between Earth and Twin Earth in Putnam's original example, to wit: the moral terms used by Earthlings rigidly designate the natural properties that causally regulate their use on Earth, whereas the twin moral terms used by Twin Earthlings rigidly designate the *distinct* natural properties that causally regulate their use on Twin Earth; hence, moral and twin moral terms *differ in meaning*, and are not intertranslatable. On the other hand, we could say instead that moral and twin moral terms do *not* differ in meaning or reference, and hence that any apparent moral disagreements that might arise between Earthlings and Twin Earthlings would be *genuine* disagreements-i.e. disagreements in moral belief and in normative moral theory, rather than disagreements in meaning.

We submit that by far the more natural mode of description...is the second...But if CSN were true, and the moral terms in question rigidly designated those natural properties that causally regulate their use, then reflection on this scenario ought to generate intuitions analogous to those generated in Putnam's original Twin Earth scenario. That is, it should seem intuitively natural to say that here we have a difference in meaning, and that Twin English “moral” terms are not translatable by English moral terms. Yet when it comes to characterizing the differences between Earthlings and twin Earthlings on this matter, the natural-seeming thing to say is that the differences involve belief and theory, not meaning” (Horgan and Timmons

1993: 130-131.)

As with Smith, here, too the characterization of the thought experiment appears to beg the question against the Descriptivist by assuming that Twin Earthlings also “use the terms ‘good’, ‘bad’, ‘right’, and ‘wrong’”. Here, too, though, it’s easy to see how to reformulate their characterization to avoid this defect:

Consider Moral Twin Earth, which is just about like good old Earth: same geography and natural surroundings...people who live in the twin United States by and large **speak a language whose terms are orthographically and phonetically just like those of English, with each term having a use quite similar to the use of it’s orthographic and phonetic counterpart in English.** Of particular importance here is the fact Moral Twin Earthlings have a vocabulary that works much like human moral vocabulary; **they use terms “G” and “B,” “R” and “W”** to evaluate actions, persons, institutions and so forth...Let us suppose that investigation into Twin English moral discourse and associated practice reveals that their uses of twin moral terms are causally regulated by certain natural properties distinct from those that...regulate **their English orthographic and phonetic counterparts.**

Reconstructing the reasoning behind this thought experiment, we can now see that it is quite like Smith’s.

1. When we focus on the similarity between our usage of “right” and their use of “R”, we see that the “most natural mode of description” of the relationship between our terms and theirs is that “moral and twin moral terms do *not* differ in meaning or reference, and hence that any apparent moral disagreements that might arise between Earthlings and Twin Earthlings would be *genuine* disagreements-i.e. disagreements in moral belief and in normative moral theory, rather than disagreements in meaning”. Call this reaction to the case ‘Judgment about Meaning’.
2. If Descriptivism were true, then the most natural mode of description of the relationship between our English terms and their Twin terms shouldn’t be Judgment about Meaning, but the judgment that our terms and theirs ‘differ in meaning and are not intertranslatable’. (Call this claim “Prediction”.)
3. So, Descriptivist theories, like CSN, are false.

Here, too, I grant #1 for the sake of argument. Our next task will be to get clear on what would need to be true in order for Prediction and Vindicate to be true. We’ll then turn to assessing the plausibility of the needed justifications for each.



## 2. Requirements on the Probative Value of Judgments about Disagreement and Meaning

What would have to be true for Judgments about Disagreement and about Meaning to have probative value for semantic theorizing? Consider first what might be said on behalf of the probative value of Judgment about Meaning. From their discussions, it's clear that Horgan and Timmons intend for the probative value of Judgment about Meaning to stem from the same source as that of the probative value of our judgments in Hilary Putnam's similar Twin Earth thought experiments. A quick defense of the claim that we should accord Judgment about Meaning probative value, then, would be an argument I'll call "Parity of Reasoning".

1. There is widespread agreement among philosophers that our judgments about whether the stuff, XYZ, on Twin Earth is water, have probative value for the purposes of identifying the correct semantics for "water" in English.
2. Moral Twin Earth, by design, is a thought experiment just like Putnam's own, except with respect to which term it targets, substituting moral terms in English for natural kind terms, like "water".
3. So, if we hold that our judgments in Twin Earth scenarios are probative with respect to the semantics for "water" in English, we should also hold that Judgment about Meaning is probative with respect to the semantics for our moral terms in English.
4. So, consistency pressures force at least most of us to treat Judgment about Meaning as probative for semantic theorizing about moral terms in English.

Notice first that this argument doesn't really vindicate the probative value of Judgment about Meaning; rather, it links its probative value to the probative value of our judgments in Putnam's Twin Earth thought experiments. To assess the full case for the former's probative value, we'll need to consider the full case for the latter's.

Backing up, then, our question is: Why should we treat our judgments about hypothetical cases like Putnam's Twin Earth as probative for the purposes of semantic theorizing? On the face of it, the claim that we should is a bit surprising. As I'll argue below, semantic theories are contingent and empirical, more like biological theories than

like paradigmatically philosophical ones, such as theories about the nature of knowledge or of normativity. While our armchair judgments about the latter may have probative value, it is unclear why our armchair judgments about hypothetical cases should have probative value for semantic theories any more than our ordinary judgments about what counts as a biological process should count as evidence for or against a biological theory.

One answer to this question is given by a familiar, Intensionalist picture of semantic competence with referential terms.

**Intensionalism:** Let “a” be the referential term under investigation. “A” must be such that

- 1) There is a set of properties,  $F$ , such that, for any environment  $E$ , what, if anything, is in the extension of “a” in  $E$  is determined by what, if anything, is  $F$  in  $E$ . In this sense,  $F$ -ness determines “a”’s *intension*.
- 2) To be competent with “a”, a speaker,  $S$ , must implicitly know that *being  $F$*  is extension-determining for “a” (at least in the sense of treating it as extension-determining). In this sense,  $S$  must implicitly know “a”’s intension.
- 3) Since implicit knowledge is a requirement on competence, competent speakers will, given a scenario (i.e. a description of an environment)  $O$  in which it is stipulated that  $b$  is  $F$ , judge that  $b$  is  $a$ .
- 4) Because the judgment *that  $b$  is  $a$  in  $O$*  is a manifestation of a speaker’s competence with “a”, it places a constraint on a semantics for “a” in that speaker’s native language.

Intensionalism can be given either a semantic or a metasemantic construal. Neo-Fregeans, such as Frank Jackson, give it a semantic construal, on which intensions are meanings. So-called ‘causal’ theories, like Putnam’s, give it a metasemantic construal, on which intensions serve as extension-fixers. This attribution is clear for Jackson’s theory, which is designed to

capture Intensionalism.<sup>13</sup> But it is perhaps less obvious that Putnam is committed to a metasemantic version, so it's worth saying a bit in defense of this attribution.

To see that Putnam (1975) is committed to something very like Intensionalism, recall his two Twin Earth thought experiments. In the first, Twin Earth is a planet very like earth, remotely located in our galaxy (1975: 139). On Twin Earth is a community stipulated to speak English, except that their word orthographically and phonetically like our term "water" applies not to H<sub>2</sub>O, but to XYZ. Putnam writes,

If a spaceship from Earth ever visits Twin Earth, then the supposition at first will be that "water" has the same meaning on Earth and on Twin Earth. This supposition will be corrected when it is discovered that "water" on Twin Earth is XYZ, and the Earthian spaceship will report somewhat as follows:

"On Twin Earth, the word 'water' means XYZ" (1975: 140).

Why does Putnam think the Earthian spaceship will make that report? Here's an important part of his answer:

The logic of natural kind terms like "water" is a complicated matter, but the following is a sketch of an answer: Suppose I point to a glass of water and say "this liquid is called water" [sic]...My 'ostensive definition' of water [sic] has the following empirical presupposition: that the body of liquid I'm pointing to bears a certain sameness relation (say, *x is the same liquid as y or x is the same<sub>L</sub> as y*) to most of the stuff I and other speakers in my linguistic community have on other occasions called "water" (1975: 141).

In this version of the experiment, the Earthians who travel to Twin Earth are moderns who know that water on earth is H<sub>2</sub>O. So, on learning that XYZ fills the lakes and oceans on Twin Earth, they refuse to recognize it as bearing the right 'same<sub>L</sub> relation' to the stuff on earth, and so refuse to recognize it as what they call "water". It's this "empirical presupposition" that's important for seeing how the Intensionalist picture fits Putnam's characterization of natural kind terms. Because of this presupposition, we moderns, as well as our space travelers, will

---

<sup>13</sup> See Jackson (1998) and (2004).

refuse to recognize XYZ as water. This judgment, together with parallel judgments about the case in which Twin Earth is in a counterfactual world (1975:148), in turn, are together to support Putnam's semantic thesis that natural kind terms are indexical and rigid over rival, Fregean theories. So, here, our disposition to call some liquid "water" or not, given appropriate characterizations of local and distant, and actual and counterfactual, environments, are to provide evidence for and against semantic hypotheses. To provide such evidence, however, Putnam's empirical presupposition must be guiding our judgments, at least implicitly.

Consider now the argument, Parity of Reasoning. According to that argument, if we hold that our judgments about XYZ have probative value for semantic theorizing about "water", we should hold that Judgment about Meaning has probative value for constructing a semantics for our moral terms, at least on grounds of consistency. Is this correct?

No. An important difference between Putnam's thought experiment and the Moral Twin Earth thought experiment is that, while the former are to trigger first-order judgments that *deploy* the targeted term ("water") and are about its extension (about water), the latter are to trigger judgments about whether disagreement is possible when two speakers use different terms in different languages and about whether those terms differ in meaning. The crucial difference is that, while the former is expressed using the targeted term to characterize its referent, Judgment about Meaning is a semantic judgment *about* the targeted term. Likewise, Smith's Judgment about Disagreement isn't a first-order judgment that deploys our moral terms, for example, a judgment about which acts are right. It's a judgment about whether the similarities in use between our terms

and theirs are disagreement-supporting. Since the notion of disagreement at issue in Smith's argument is disagreement about truth, not in attitude, Judgment about Disagreement is, or at least presupposes, a semantic judgment, namely, that the similarities in use provide sufficient similarity in meaning to support the expression of disagreement using the targeted terms.

As we'll see in the next section, this difference is important for assessing the probative value of both judgments. For now, though, notice that this difference means that, while the truth of Intensionalism would provide good reason to treat our judgments about Putnam's scenarios as probative, it would not provide good reason to treat either Judgment about Meaning or Judgment about Disagreement as probative. To see this, recall that Intensionalism is a thesis about what is required to competently use a term. Competent use of "a" requires that one's use of "a" is guided by one's implicit understanding that being *a* requires having *F*. This means that when a speaker known to be competent with "a" says "b is a" upon appreciating that *b* is *F*, that judgment provides evidence for a semantic theory that holds that "a" is guided by an appreciation of something's being *F*. Likewise, when a competent speaker, as in Putnam's thought experiments, refuses to call *b* "a", this provides evidence that being *F* isn't guiding such usage. The reason why the guidance such uses are to manifest puts a constraint on a semantic theory, according to Intensionalism, is because what's doing the guiding is implicit knowledge of an extension-fixer, either by telling us directly about the semantics (as on Jackson's theory) or the metasemantics (as on Putnam's).

Suppose Intensionalism is true and to be competent, a speaker *S* of *L* must have implicit knowledge of the extension-fixers for *L*'s terms. Such knowledge would not by

itself confer the further knowledge of which similarities in use between  $L$ 's terms and those of a distinct, hypothetical language  $L'$ , would make for sameness of semantic significance. Nor would it by itself confer the further knowledge of which such similarities sufficed for the possibility of cross-linguistic disagreement about truth. But for the truth of Intensionalism alone to make Judgments about Meaning and Disagreement probative, the knowledge Intensionalism requires of a competent speaker  $S$  would need to suffice for possessing these further pieces of knowledge. The third premise in Parity of Reasoning is false, then, and those who regard our judgments in Putnam's thought experiments as probative are under no consistency pressure to treat Judgment about Meaning or about Disagreement as probative as well.

For judgments like Judgment about Meaning and Judgment about Disagreement to be probative, then, we need a picture of semantic competence stronger than Intensionalism, something like:

*Semantic Intensionalism:* Competence with our moral terms in English requires knowledge of which cross-linguistic similarities in use between our terms and those of any rival, hypothetical language,  $L'$ , make for sameness in meaning and so the possibility of using our moral terms to express cross-linguistic disagreement with speakers of  $L'$ .

If Semantic Intensionalism were true, Judgment about Meaning and Judgment about Disagreement would be probative for semantic theorizing about moral terms in English. Its truth, though, isn't self-evident. Treating those judgments as probative requires either some reason to think it is true or some other explanation of their evidential relevance. At this point, we might think the work of a defender of Descriptivism is done: Having provided room for rational doubt about their probative value, the burden of proof is shifted back onto those who treat them as such.

More, though, may be said in defense of Descriptivism. *Prima facie*, Semantic Intensionalism is implausible. If it were true, its truth would stem from some feature special to our moral vocabulary as no one, I take it, would defend a parallel claim about our ordinary, empirical terms. To see this, consider an ordinary English speaker who is competent with the term “elm”. Such a speaker is able to use the term to communicate and coordinate with others who are similarly competent. If we accept Semantic Intensionalism for “elm”, though, we are forced to accept one of two implausible consequences. First, we could hold that that ability of a monolingual English speaker to communicate with “elm” comes with it an ability to distinguish which cross-linguistic similarities make for sameness of meaning for any hypothetical language L’ (independently of knowing the facts about which properties causally regulate the use either of “elm” or of terms in L’). Or, second, we could hold that failure to possess the latter ability shows that speakers who are able to communicate and coordinate using “elm” are nonetheless not competent with that term. The first of these commitments, I take to be *prima facie* extremely implausible. To see that the second claim about what semantic competence requires is not only surprising, but also an implausible, we will need to think about which notion of semantic competence is appropriate for semantic theorizing. In the next section, I consider three strong reasons to doubt Semantic Intensionalism and, along the way, address this issue about semantic competence.

### **3. Assessing Semantic Intensionalism**

#### *Semantic and Metasemantic Theories are Contingent and Empirical*

Before addressing the question of which notion of semantic competence is most appropriate for the purposes for semantic theorizing, it will help first to step back and

remember what a semantic theory is and what it is supposed to do. A *semantic theory for a language, L*, assigns meanings to each of *L*'s simple expressions and identifies rules of composition for complex expressions such that the meanings of the latter are the products of the meanings of the simples out of which they are composed, together with the rules they exemplify. Such a theory should fit with and help explain data about what competent speakers are able to do with *L*, centrally, communicate, coordinate, and collect information.<sup>14</sup>

This observation helps illuminate which notion of semantic competence is relevant for semantic theorizing. As just noted, the phenomena a semantic theory's meaning assignments are to help explain are the ability of speakers of *L* to communicate, coordinate with other speakers, and collect information using *L*. Among the uncontroversial data for semantic theorizing are ordinary uses, as well as felicity judgments about such uses, by speakers who have this ability. To be a competent speaker of *L*, in the relevant sense, then, just is to have this ability to coordinate, communicate and collect information using *L*. Notice, though, that this is an ability humans acquire for good chunks of their native language at quite early ages, prior to any formal education, acquired ability for careful reflection, or extensive world experience (Bloom 2002). Whatever is required for a person to have this ability, it cannot be more demanding than would explain this fact about language learning.

Earlier I noted that Intensionalism, as a theory of semantic competence, may be given a semantic construal or a metasemantic construal, so a quick reminder of what a metasemantic theory in the relevant sense is would be helpful here. A *metasemantic*

---

<sup>14</sup> See also Yalcin (manuscript).



*theory for L* is a theory that tells us why and how it is that the phonetic and orthographic sequences that correspond to the expressions of *L* have the meanings our best semantic theory represents them as having.

How do these two kinds of theory relate to one another? Start with what we know:

- 1) There are noises we able to make, scribbles we are able to produce. *Somehow*, some of these noises and scribbles, those that correspond to words in the English language, acquired the meanings our best semantic theory, *T*, assigns to them.
- 2) Using a language with those meanings allows competent speakers to communicate, coordinate, and collect information.

Keeping this in mind, what should we think about the relationship between our theories?

Should we think: Lucky are we! For we narrowly missed living in a world in which we spoke a language, but one which didn't allow us to coordinate our activities or communicate and collect information. Or, rather, should we think: It's no accident! Part of the explanation for why our expressions have the meanings that they do is that, in using expressions with *those* meanings, we are able to communicate, coordinate, and collect information. The latter is clearly so much more plausible than the former, that I propose the following,

**Constraint:** A metasemantic theory should help us understand how the sequences that make up *L* acquired semantic significance such that speakers of *L* are able to do what they do with *L*'s expressions, centrally, communicate, coordinate, and collect information.

What this means is that semantic notions are theoretical notions that earn their keep by figuring in the best explanations of the phenomena it is their job to explain. Moreover, since the phenomena to be explained—speakers' use of phonetic and orthographic sequences to promote their aims—are contingent, empirical features of the actual world, metasemantic and semantic theories are themselves contingent and empirical; had the actual world differed in relevant respects, the ability of speakers to use a language to

coordinate, and what explained that ability to coordinate, would itself have been different. The features of our world that support communication and coordination with natural languages are features that distinguish our world from others, whether considered as actual or counterfactual. This suffices to guarantee that those features are contingent and empirical.<sup>15</sup>

We can begin to consider the question of what types of consideration support semantic theories, then, by considering how contingent and empirical theories, e.g. psychological or biological theories, are justified. The usual method for constructing and justifying such theories is scientific; we identify an array of phenomena to be explained and begin theory construction by developing a technical vocabulary that earns its keep by its ability to figure in empirically well-confirmed explanations of those phenomena. The plausibility of proposed explanations of specific phenomena are measured against their fit both with data and total theory; explanations that merely fit with local observations without clear fit with what's known overall about a subject matter are ad hoc.<sup>16</sup>

In these respects, semantic and metasemantic theories are like biological ones. This also suggests that, as with the notion of a process for biological theorizing, the notion of meaning for the purposes of semantic theorizing is a theoretical notion; the work meanings are to do in

---

<sup>15</sup> A comparison to physicalism about the mental might be helpful here. While a physicalist about the mental is committed to the thesis that every world that is a minimal physical duplicate of the actual world is a mental duplicate, she is not committed to saying that the physical features that ground the mental properties instantiated at the actual world ground them at every world, whether considered as actual or counterfactual. Identifying which physical properties are plausible candidate grounds for our mental properties at the actual world is properly the job of neuroscientists. Analogously, it is properly the job of semanticists to identify what natural language signs mean at the actual world and the job of the metasemanticist to identify which contingent features of the actual world are plausible candidate grounds for those meanings.

<sup>16</sup> See Hempel (1966) for discussion.

explaining communication and coordination constrains how that notion is best understood.<sup>17</sup> Which such notion is explanatorily best, in turn, constrains a semantic theory's particular meaning assignments. Together, these facts about how other contingent, empirical theories are justified and about what constrains a plausible semantic theory's meaning assignments provide our first grounds for doubt about the probative value of Judgment about Meaning and Judgment about Disagreement: Whether or not the correct semantic theory for English would assign the same meaning to "...is right" as would a hypothetical semantic theory for a hypothetical language, stipulated to contain a term, "...is R" whose use (though not extension) is like that of "...is right" is settled both by which notion of meaning, of the many candidates, does the best overall explanatory job, and which meaning-assignment for "...is right" does the best particular one. It would be quite surprising if the cross-linguistic judgments about sameness and difference of meaning of ordinary, monolingual speakers were in general a good guide to which such judgments are most explanatorily useful. Again, we might usefully compare this case to the case of biological explanation: It would be a surprise if non-expert judgments about when two token biological processes were of the same or different type were probative, given that the notion of a biological process is a theoretical one whose choice depends in part on its explanatory usefulness.

### *Language Learning*

These first grounds for doubting Semantic Intensionalism are not grounds for doubting Intensionalism. What's known about language learning provides our second grounds for doubting the former, as well as ground for doubting the latter. We saw above that, for the purposes appropriate for semantic and metasemantic theorizing, semantic competence is an

---

<sup>17</sup> See also Yalcin (forthcoming).

ability to communicate, coordinate, and collect information. By the end of high school, the average child has learned sixty-thousand words, roughly the rate of ten per day. Many of these are learned from only one encounter, from linguistic contexts alone, or from contexts with impoverished contextual cues (Bloom 2002 chp.3). In one type of well-replicated experiment, children are on one occasion shown two objects, one familiar and one unfamiliar. Afterwards, they reliably select the unfamiliar object, when asked “where’s the [new word]?” (Bloom 2002: 66) Here, the only property distinctive of the new word’s referent is its unfamiliarity, a property it has only on this introductory occasion of use. In another type of study, children correctly learn to use a word to pick out an object in a context in which the object is not the most perceptually salient one. (The word is introduced while they are holding and looking at a toy in their hands, while the object referred to is out of their sight, in a bucket. (Bloom 2002: 6304)) In many cultures, language is not taught, but learned from overheard speech. Blind children learn words at the same rates as sighted children. (Bloom 2002: 59)

These facts about language-learning don’t fit well with either Intensionalism or Semantic Intensionalism. Intensionalism requires that children become competent with a term by acquiring an ability to linguistically track a term’s referent by learning to associate with it a set of features something has to have to be in that referent in *any* environment. But, given the impoverished conditions in which children learn language, this is highly implausible. Given the range of conditions under which different children learn the same word, it is even less likely that, for each word, each child is acquiring the same such method. A blind child who is able to track water by its characteristic taste and feel and for whom water is typically found in bathtubs and sippy cups will have a different method from that of a sighted child who lives by the ocean. Each child is able to use “water” to communicate and coordinate with other

speakers of English, even though neither may be a reliable tracker using the other's method in her different environment. Despite failing to possess the type of competence Intensionalism requires, each has the ability that it is the job of a semantic theory to help explain.

As implausible as these facts make Intensionalism as an account of the type of semantic competence relevant for semantic theorizing, they make Semantic Intensionalism still less plausible. Semantic Intensionalism requires that, in order to be competent with moral terms, a speaker must be able to tell, from stipulated similarities in use, whether she is able to disagree with speakers of a hypothetical language with which she is not competent. Given that moral vocabulary is among the vocabulary acquired at an early age (Kagan 2007), it seems implausible that this is so. To suppose otherwise is, in effect, to hold that in order to use "...is right" and "...is wrong" to coordinate with others in English, a child must acquire the skills of a field linguist engaged in radical translation.

The data about language learning suggests that it is much more plausible that, in acquiring competence with a term, each child acquires some one or other highly accidental, even highly fallible, method for tracking its extension in the limited range of environments in which she typically finds herself. As we'll now see, such an account of semantic competence fits very well with Millikan's fully externalist metasemantic rival to Intensionalism.

### *Millikan's Metasemantics*

What's known about language learning provides good reason to doubt both Intensionalism and Semantic Intensionalism. It also provides good reason to accept Millikan's anti-Intensionalist rival metasemantics. Millikan defends a fully externalist account of how it is that empirical terms have the semantic significance they do. Descriptivism, in the sense used here, is the view that moral terms share a semantics with ordinary empirical terms. So, it is

open to the Descriptivist to adopt Millikan's theory as a theory about what significance moral terms have and how they come to have it. Millikan's theory, however, accords our judgments about thought experiments like Putnam's no probative value. Seeing why this is so and why her theory is independently plausible will help provide us with our third grounds for rejecting Semantic Intensionalism. It should be noted, though, that while Millikan's view does enjoy independent plausibility, as we shall see, our third grounds for rejecting Semantic Intensionalism do not depend upon the claim that Millikan's account is wholly correct. What the anti-Descriptivist needs is that *no* metasemantic theory for empirical terms incompatible with Semantic Intensionalism is correct. Thus, Millikan's view provides but one option for the Descriptivist.

Millikan accepts that semantic and metasemantic theories are contingent, empirical theories that receive their support in the same way other such theories do, by their ability to explain the phenomena it is their point to explain. In the case of a semantic theory for a language *L*, the phenomena centrally to be explained are how, using *L*, competent speakers are able to communicate, coordinate, and collect information. The job of a metasemantic theory, for Millikan, is to meet Constraint, that is, to help us understand how it is that the terms in *L* acquired the semantic significance our best semantic theory for *L* assigns to them, such that, using terms with those significances, competent speakers are able to communicate, coordinate, and collect information. [Millikan (1987).]

An unfortunate consequence of this original approach to metasemantic theorizing is that her theory is stated using technical terms of her own coinage, many of them interdefined.<sup>18</sup> In

---

<sup>18</sup> This approach is highly original to the philosophical literature on metasemantics. Of course, an approach to semantic theorizing driven by empirical data has long been the practice of linguists.

addition, part of the plausibility of her theory lies in the careful development of its details.

These together make a brief characterization of her theory difficult. Below I'll do my best to put enough of her view on the table, using as few of her technical terms as possible, to put us in a position to assess the replies to our challenges to Descriptivism about moral terms her theory makes available.

For Millikan, an expression, *E*, in a language, *L*, has at most, two components to its meaning, a stabilizing function and an extension. Since the Descriptivist about moral terms holds that such terms function semantically in just the way ordinary, descriptive terms do, my focus here is on her account of empirical terms, which have both.

First,

“...the function—I call it a ‘stabilizing function’—of a conventional language form is roughly its survival value. It is an effect it has had that encouraged speakers to keep reproducing it and hearers to keep responding to it in a roughly uniform way, each relying on the settled dispositions of the others. For example, the sentence forms that are labeled ‘indicative’ in a language have as one stabilizing function to produce in hearers true beliefs, their content varying with the words arranged into this form. Indicative forms survive, in large part, because the imparting of true beliefs is often of interest to speakers and hearers alike... meaning is, simply, stabilizing function.” (2010: 53. For details and defense, see her 1987.)

It's important here to avoid an unfortunately common misunderstanding of Millikan's notion of a proper function, of which stabilizing functions are a type. It is often mistakenly thought that, for Millikan, proper functions are biological functions, the product natural selection. As she writes, though,

'proper functions' [are] 'biological' *in a broadly metaphorical sense*, derived from perceptual tuning, from various kinds of learning or from cultural selection as well, of course, as from natural selection. (2012: 104, italics mine.)

So, stabilizing functions are not biological functions, though they are like biological functions in being selected for by their usefulness either to the objects that have those

functions or to the producers and consumers of such objects.

Second,

“Besides this most basic dimension of meaning, I have proposed that a secondary kind of meaning for many terms and phrases concerns items in the world to which they necessarily correspond when managing to serve their stabilizing functions in accordance with historically normal explanations for this achievement. That is, many terms and phrases have referents or extensions in conjunction with their stabilizing functions.” (2010: 54. For details and defense, see her 1987.)

“Normal explanation” is a technical term for Millikan; crucially, “normal” does *not* mean “typical” or “usual” (1987: 5). Rather, a normal explanation is, roughly, the explanation for how an item, e.g. an expression, has performed its function historically, such that, it is because ancestors of current individual such items have successfully performed their function under the conditions mentioned in such explanations that explains why those current individuals exist and that tokens of the type will continue to be produced. (1987: 33-34.) The extension of a term is that onto which the tokening of a term *normally* maps: When conditions are normal, the tokening of an indicative sentence containing an empirical term (roughly) co-varies with a member of its extension.

Millikan holds that the mechanisms that explain human speech and its intentionality are of the same kind as those that explain more simple cases of communication between animals. So, these ideas may perhaps be most simply illustrated with a non-human example. (A bit of background: A “consumer” of a representation is (roughly) any object that makes use of representations as representations in order to perform one of its functions.)

“...consider beavers, who splash the water smartly with their tails to signal danger. This instinctive behavior has the function of causing other beavers to take cover. The splash means danger, because only when it corresponds to danger does the instinctive response to the splash on the part of the...[other] beavers, the consumers, serve a purpose. If there is no danger present, ... [the consumers] interrupt their activities uselessly.



Hence, that the splash corresponds to danger is a normal condition for the proper functioning of the...[consumer] beavers' instinctive reaction to the splash. (It does not follow, of course, that it is a usual condition. Beavers being skittish, most beaver splashes...occur in response to things not in fact endangering the beaver.) In the beaver-splash semantic system, the time and place of the splash varies with, "corresponds to" the time and the place of the danger. The representation is articulate: properly speaking, it is not a splash, but a splash-at-a-time-and-a-place. Other representations in the same system, splashes at other times and places, indicate other danger locations." (1989: 288.)

Here beaver tail splashes are representational devices with a stabilizing function to coordinate producer uses with consumer responses by co-varying at a time and a place with beaver-danger at that time and place.<sup>19</sup> Given this Normal co-variance, beaver-danger at the time and the place of the splash is what such splashes represent (albeit often mistakenly).<sup>20</sup>

What would be required for a speaker to be competent with a language whose expressions had stabilizing functions as their meanings? Although Millikan herself doesn't offer one, there's an account of semantic competence, at least for referential terms, naturally suggested by her work. Here is one such suggestive passage:

"...agreement in methods [for identifying a term T's extension] often is not necessary to determining agreement in judgments [expressed using T], hence not for communication, nor for acquiring stable knowledge. Moreover, as a practical matter, their own agreement in judgment with others is often the only thing actually discerned by language learners and users as a check on their usage, hence the only factor (of this kind) controlling proliferation of an extensional term's tokens. So it cannot be part of such a term's stabilizing function to implant intentional attitudes towards its extension or members of its extension (towards X or towards Xs) as recognized in any *particular* way, or as thought of under any *particular* description. Our basic extensional terms do not have handed-down conventional intensions." (2010: 57, my emphases.)

---

<sup>19</sup>For more on stabilizing functions, see 1987: 3-4 and 31-32.

<sup>20</sup>This is a bit rough. Depending upon how the facts about the beavers' environment get filled out, it may be that splashes map onto a more specific source of beaver-danger, e.g. predatory birds at a time and a place.

On the suggested picture, a speaker,  $S$ 's, being competent with an empirical term,  $T$ , requires  $S$ 's having some method, or, more likely, methods, for tracking  $T$ 's extension in her environment often enough that she is able to communicate, coordinate, and collect information with other, similarly competent speakers. The method she deploys needn't be one that allows her to track  $T$ 's extension in any environment in which she might actually find herself, let alone any possible environment. The features by which she tracks  $T$ 's extension in an environment may not only be highly accidental, they may be neither necessary nor sufficient for being in  $T$ 's extension. This feature of her view will be important for thinking about why, on that view, neither Judgment about Meaning nor Judgment about Disagreement have probative value, so it is perhaps helpful to have a somewhat concrete illustration. Think again about a young child's ability to identify and request water using "water". In the child's experience, water has always been found in bathtubs and sippy cups. Suppose again that the child is blind; she identifies water by a characteristic feel and taste. Given her environment, tracking these features has allowed her to track water often enough that continued use of "water" has a point. But put her in another environment, in which water is found in glasses, but not tubs and sippy cups, or in which another liquid has the same feel and taste, and her methods will no longer be reliable. That non-reliability in other environments, however, does not undermine her ability to use "water" to track water often enough that there is a point to her using the term. This is all that is required for her uses *in her standard environment* to count as competent ones in the sense of competence of interest to semanticists, i.e. as providing data for semantic theorizing.

One advantage of Millikan's metasemantic view is that it straightforwardly meets Constraint. The explanation for why empirical terms have the extensions they do is that Normally mapping onto those extensions allows producers and consumers of such terms to

fulfill the purposes of their tokenings. Second, unlike Intensionalism, the picture of semantic competence her theory suggests fits with what we know about primary language learning. What should we conclude from this? One reasonable conclusion is that this data about language learning is insufficient to warrant acceptance of any positive view about the nature of human semantic competence. That said, it must be acknowledged that it provides some support for Millikan's account and some evidence against both Intensionalism and Semantic Intensionalism. Suppose we conclude that, insofar as we are going to go in for forming positive judgments about what semantic competence requires of speakers and about what determines the extensions of our terms, we do best to conform our judgments to Millikan's theory. Our question, then, is what should we then conclude about the probative value of Judgments about Disagreement and Meaning and similar judgments? Answering this question will involve some work. Fortunately, we can make a start by noting the implications of accepting her theory for Putnam's Twin Earth thought experiments, about which she is explicit:

Recall Putnam's fable about Twin Earth, which is identical in all ways to Earth except that what they call 'water' on Twin Earth has the molecular structure XYZ. XYZ resembles water in every surface respect by which water was identified on Earth in 1750, before modern chemistry determined water to be H<sub>2</sub>O. On [the theory I have suggested], as on Putnam's own, XYZ was not within the extension of our word 'water' even in 1750. The stabilizing function, the current meaning, of a word rests on what has, as a contingent matter of fact, been holding its usage in place, effecting agreement among users and for users with themselves, despite the use of a variety of alternative recognition techniques. Dispositions that people might have to apply a term in new circumstances that would accidentally continue to support agreement but for new reasons are not relevant to its current meaning. On Earth in 1750, the deep structure XYZ had borne no responsibility for stabilizing the earth...term 'water'. It was not because XYZ could manifest itself in the many alternative characteristic ways by which various people at various times had been identifying water...that they had managed to agree...in judgments, but because H<sub>2</sub>O manifested itself in these various ways...

Notice that this argument for Putnam's result does not depend, as his did, on some special linguistic rules we have all learned determining that 'water' must always name 'the same liquid' and legislating some common way that we should all understand 'same liquid'. The reference of 'water' is entirely *direct*. It has no defining intension. (2010: 65)

This passage makes clear why her view accords our first-order judgments about XYZ no probative value. But it also helps make clear why her view accords no such value to Judgments about Meaning and Disagreement or to any similar semantic judgment about hypothetical cases. On that view, the notions of stabilizing function and of reference are the semantic notions that do the best job of explaining language use. So, whether a term,  $T_1$ , in a language  $L_1$ , shares a semantics with another term,  $T_2$ , in a distinct language,  $L_2$ , depends upon whether or not  $T_1$  and  $T_2$  share a stabilizing function and (when  $T_1$  and  $T_2$  are empirical terms) a referent. Knowing that this is what we'd need to know to know whether these terms are intertranslatable is empirical knowledge not itself required for semantic competence, on the picture Millikan's theory suggests. On that picture (and granting the Descriptivist for the sake of argument, as Smith and Horgan and Timmons intend, that moral terms are ordinary, descriptive terms<sup>21</sup>), to be competent with terms like "right", one need only have *some* method for identifying what's right *often enough* that one is able to collect information about what's right and communicate that information with others.

Now first consider Smith's Judgment about Disagreement. Millikan's picture of semantic competence allows that a speaker may be competent with our moral terms compatible with feeling, mistakenly, that, Smith's description of his "rather different

---

<sup>21</sup> To be clear, Millikan is not herself a Descriptivist about moral terms. The strategy here is to block the anti-Descriptivist's conclusion by using her metaseantics for empirical terms to provide an example of a semantics for such terms, available to the Descriptivist, on which Judgments about Meaning and Disagreement have no probative value. For Millikan's account of such vocabulary, see her 2005: 166-86.

community” suffices to settle whether or not disagreement is possible with that community. Such a speaker might then judge that disagreement is possible, though she lacks the information necessary to determine whether it is. What this means is that competent speakers’ judgments about the possibility of disagreement with hypothetical, rival speech communities are not a guide to the possibility of disagreement, as what would secure that possibility is wholly external to the heads of ordinary competent speakers. But if that’s so, then Judgment about Disagreement has no probative value for semantic theorizing and Vindicate is false.

Similar considerations apply to the Horgan and Timmons Moral Twin Earth scenario. It may be that the truth of CSN, were it modeled on Putman’s account of kind terms, would generate the prediction that, in that stipulated scenario, our moral terms and those of our Twins have different meanings. And it may be that, combined with Intensionalism, CSN generates the prediction that competent speakers of English will refuse to call “right” at least some of what the Twearthians are stipulated to call “R”. That does not mean, however, that the truth of Descriptivism generates these predictions. A Descriptivist, proponent of CSN or otherwise, might instead propose to model their semantics and metasemantics for moral terms on Millikan’s metasemantics for empirical terms. If she does, she would do well to begin by noting that Millikan’s fully externalist theory makes *no* predictions about what intuitions to expect competent speakers to have in Moral Twin Earth cases, predictions neither about what first-order judgments speakers will have, nor about judgments about the possibility of disagreement with Moral Twin earthlings. So, Prediction is false.

Moreover, were competent speakers to have judgments of either kind, the account of semantic competence her metasemantics suggests would not accord them probative value. On that view, recall, a speaker *S*'s competence with an empirical term *T* requires that *S* has some method for tracking *T*'s extension often enough in the environments in which she typically finds herself that continuing to use *T* serves some of her purposes, especially the purposes of communicating and coordinating with others and collecting information. The methods she has for tracking a moral term's extension may be idiosyncratic and highly dependent upon features specific to those environments. Ask *S* to use *T* to characterize a different, twin environment and she will deploy a method, reliable enough here, that may well mislead there.

What determines whether the judgments about those different environments expressed using English moral terms are mistaken or not are the facts about what those terms mean. On a fully externalist, metasemantics, such as Millikan's, those meaning-determining facts are independent of rules of thumb we rely on to arrive at the judgments we express using moral terms. Whether we can disagree or not with Moral Twin Earthians is a matter of what we and they mean, whether there are similarities in meaning sufficient to support disagreement. Whether that is so, on an externalist view, depends at least on whether we are tracking the same properties. When it is stipulated that we're not, it's stipulated that the similarities in meaning are not disagreement-supporting. This may be so even if competent speakers have the sense that they are. This sense of disagreement, then, would be no guide to the genuine possibility of disagreement. Given this, to treat either Judgment about Disagreement or Judgment about Meaning as probative for the purposes of semantic theorizing about moral terms is to beg the question

against a fully externalist account of those terms, like one modeled on Millikan's metasemantics for empirical terms.

## **Conclusion**

What's central to Descriptivism, in our sense, is the idea that moral terms in English share a type of semantics with ordinary, descriptive terms. Such a view might get developed in any number of ways, filling out what the appropriate semantics would need to be like with a favored semantics for ordinary descriptive terms. One model might be the Putnam doctrine on natural kind terms. But another might be given by a more fully externalist view, such as Millikan's. Insofar as Putnam's view is Intensionalist, it is perhaps not best suited for the Descriptivist's purposes. As we've seen, Intensionalism by itself doesn't suffice to leave the Descriptivist vulnerable to the Smith or the Horgan and Timmons objections, as the probative value of our judgments in those cases (assuming their proponents are correct in the judgments ascribed to us) require something stronger than Intensionalism, something more like Semantic Intensionalism. Perhaps, though, there is a way to devise a new thought experiment involving a hypothetical speech community that would leave a Putnamian Descriptivist open to falsified prediction. It is open to the Descriptivist, in that case, to model her metasemantics for moral terms on Millikan's for empirical terms. As we've seen, if she does so, she'll have principled grounds for rejecting the probative value of our judgments in those cases.

Indeed, the Descriptivist may claim something stronger. Millikan's metasemantics and account of semantic competence are not merely possibly correct; they enjoy some empirical support. Moreover, the supporting evidence tells against both Intensionalism and Semantic Intensionalism. One perfectly sensible response to the data is agnosticism

about metasemantics and semantic competence; instead, we should await the arrival of more and better data before accepting any, even working, hypotheses about either. Agnosticism, though, does not help the fan of our thought experiments. If we are wholly in the dark about what might determine semantic facts and about what semantic competence consists in, we should not assume that our judgments about the possibility of disagreement with rival, hypothetical speech communities have any probative value for semantic theorizing.

A second, sensible response to the data on language-learning would be to treat it as providing sufficient evidence to accept Millikan's view as a working hypothesis for the assessment of claims about which judgments of competent speakers, first-order and otherwise, plausibly constrain semantics. As these are our two plausible responses, and neither of them supports their probative value, we are left with the conclusion that we should accord our intuitions in these cases no probative value for the development of a semantics for moral terms.

## References

- Bjornsson, Gunnar and Tristram McPherson Forthcoming. "Moral Attitudes for Non-cognitivists", *Mind*.
- Bloom, Paul 2002. *How Children Learn the Meanings of Words*. Bradford Books.
- Boyd, Richard 1988. "How to be a Moral Realist", *Moral Realism*, ed. G. Sayre-McCord. Ithaca: Cornell University Press, pp.
- Brink, David 1989. *Moral Realism and the Foundations of Ethics*. New York: Cambridge University Press.
- Copp, David 2000. "Milk, Honey, and the Good Life on Moral Twin Earth", *Synthese* (124) pp. 113-137.
- Hare, R.M. 1952. *The Language of Morals*. Oxford: Oxford University Press.
- Hempel, Carl 1966. *Philosophy of Science*. Edgewood Cliffs: Prentice-Hall.



- Horgan, Terrence and Mark Timmons 1993. "New Wave Moral Realism Meets Moral Twin Earth", in *Rationality, Morality, and Self-Interest*, in ed. John Heil, Rowman and Littlefield, pp.115-133.
- Jackson, Frank 1998. *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Jackson, Frank 2004. "Why We Need A-intensions", *Philosophical Studies* (118) pp.257-277.
- Kagan, Jerome 2007. "Morality and its Development", in *Moral Psychology: The Neuroscience of Morality*, ed. Walter Sinnott-Armstrong, Cambridge: MIT Press, pp. 297-312.
- Laurence, Stephen, Eric Margolis, and Angus Dawson 1999. "Moral Realism and Twin Earth", *Facta Philosophia* (1), pp. 135-165.
- Merli, David 2002. "Return to Moral Twin Earth", *Canadian Journal of Philosophy* (32), pp. 207-240.
- Millikan, Ruth 1987. *Language, Thought, and Other Biological Categories*. Cambridge: MIT Press.
- Millikan, Ruth 1989. "Biosemantics", *Journal of Philosophy* (86), pp.281-297.
- Millikan, Ruth 2005. *Language: A Biological Model*. Oxford: Oxford University Press.
- Millikan, Ruth 2010. "On Knowing the Meaning; With a Coda on Swampman", *Mind* (119) pp.43-81.
- Millikan, Ruth 2012. "Reply to Rescorla", in *Millikan and her Critics*, ed.s Ryder, Dan; Justine Kingsbury, and Kenneth Williford.
- Plunkett, David and Tim Sundell 2013. "Disagreement and the Semantics of Normative and Evaluative Terms", *Philosophers' Imprint* (13) pp.1-37.
- Putnam, Hilary 1975. "The Meaning of 'Meaning'", *Language, Mind, and Knowledge. Minnesota Studies in Philosophy of Science* (7), pp. 131-193.
- Ridge, Michael 2014. *Impassioned Belief*. Oxford: Oxford University Press.
- Rubin, Michael 2008. "Sound Intuitions on Moral Twin Earth", *Philosophical Studies* (139) pp. 307-327.
- Ryder, Dan; Justine Kingsbury, and Kenneth Williford 2012. *Millikan and Her Critics*. Basil Blackwell.
- Sayre-McCord 1997. "'Good' on Moral Twin Earth", *Philosophical Issues* (8), pp. 267-292.

Smith, Michael 1994. *The Moral Problem*. Wiley-Blackwell.

van Roojen, Mark 2005. "Knowing Enough to Disagree: A New Response to Moral Twin Earth", *Oxford Studies in Metaethics* (1), ed. Russ Shafer-Landau. Oxford: Oxford University Press, pp. 161-193.

Wedgwood, Ralph 2001. "Conceptual Role Semantics for Moral Terms", *Philosophical Review* (110) pp.1-30.

Wedgwood, Ralph 2007. *The Nature of Normativity*. Oxford: Clarendon Press.

Yalcin, Seth. Manuscript. "Semantics and Metasemantics in the Context of Generative Grammar", <http://yalcin.cc/resources/yalcin.2013.meta.pdf>