

---

HEIMIR GEIRSSON

## MORAL TWIN-EARTH AND SEMANTIC MORAL REALISM

**ABSTRACT.** Mark Timmons and Terry Horgan have argued that the new moral realism, which rests on the causal theory of reference, is untenable. While I do agree that the new moral realism is untenable, I do not think that Timmons and Horgan have succeeded in showing that it is. I will lay out the case for new moral realism and Horgan and Timmons' argument against it, and then argue that their argument fails. Further, I will discuss Boyd's semantic theory as well as attempts to improve upon it, raise serious problems for these semantic accounts, and suggest an alternative view that accounts for our use of moral terms.

### 1. CAUSAL THEORIES OF REFERENCE

In the early 1970's a number of philosophers advanced arguments against the prevailing account of the meaning of proper names, the description theory of proper names. According to the description theory, advocated for example by Russell, Frege, Wittgenstein, and Searle, proper names had meanings that could be identified with either a definite description (Russell and Frege) or with a cluster of definite descriptions (Wittgenstein and Searle). Furthermore, the reference of the names was determined by their meanings. If, for example, "Thales" means "the Greek philosopher who held that all is water", then "Thales" refers to whoever fits the identifying description; i.e., to whoever was the Greek philosopher who held that all is water. A consequence of this account is that it is a necessary truth that Thales was the Greek philosopher who held that all is water.

Against the description theory Kripke, Donnellan, Barcan-Marcus, Putnam, and Kaplan, to name a few, advanced the causal theory of reference. According to the causal theory of reference a name is attached to an object via an initial act of baptism, and the object is the semantic meaning of the name. The name might then be passed from one language user to the next. As long as the language users intend to use the name with the same reference it had when passed to them, the name maintains its reference. As an example, there is a causal chain

that connects our use of “Thales” to the initial baptism of Thales, thus maintaining its reference. Whether or not the person at the tail end of the chain held that all is water is inconsequential for the reference of the name to succeed. Instead of referring to whatever person held that everything is water, as a description theory would have it, the name refers to Thales, the person at the end of the causal chain. The name refers via this causal chain and not via any properties the object might otherwise have, i.e., the name directly designates Thales. In this way the causal theory accounts nicely for our intuition that it is contingently true that Thales believed all is water.

Another liability of description theories is that according to them true identity statements are only contingently true. The causal theory accounts for how true identity statements are necessarily true. One element of the causal theory of names is that proper names are rigid designators, meaning that they refer to the same objects in all possible worlds in which the objects exist, and this feature helps explain how it is that true identity statements, such as “Samuel Clemens is Mark Twain”, are necessarily true. If “Samuel Clemens” and “Mark Twain” are rigid designators, then they designate the same objects in all possible worlds. Given that, there will be no worlds in which Clemens is not Twain, so it is a necessary truth that Clemens is Twain.

The causal theory of reference has also been used to argue that natural kinds have essential properties. Suppose I decide to call a type of liquid “water”, and then find out that this type of liquid has an atomic structure  $H_2O$ . Then, since “water” is a rigid designator that names the same type of material in all possible worlds in which it exists, “water” refers to water in all possible worlds (in which water exists). It might be the case that water has different phenomenal properties in different worlds. It might, for example, be green in some worlds. But, just as it was contingently true that Thales held that all is water, it is contingently true that water has the phenomenal properties it has, such as being a clear liquid in this world and a green liquid in some other possible world. What makes water water is its atomic structure, not its phenomenal qualities. And since water has the atomic structure  $H_2O$ , it has that atomic structure in all possible worlds. It is a necessary, or an essential property of water to be  $H_2O$ . Nevertheless, it would be a mistake to say that it is true in virtue of the meaning of the words that water is  $H_2O$ , for it took substantial research to uncover the fact that water is  $H_2O$ .

Hilary Putnam put the causal theory of reference to a different task. He argued that our thoughts about external objects cannot be

fully individuated by what is in our heads. Putnam construed a number of thought experiments to support this view. Suppose, for example, that we have two persons, Oscar and Toscar, who are exact duplicates of each other and who live in worlds that are exact duplicates, except for one thing. Oscar lives on earth where there is water while Toscar lives on twin-earth where there is no water, but where there is substance that looks like and feels like water and, in fact, has all the phenomenal properties of water. But this substance on twin-earth is not H<sub>2</sub>O. Instead it is an unknown substance, XYZ. Everything else is the same on earth and twin-earth. Now, imagine that Oscar has a thought about water, namely that it is quenching. At the same time Toscar has a thought about twater, namely that it is quenching. In Putnam's view, Oscar and Toscar have different thoughts, for Oscar's thought is about water while Toscar's thought is about twater. The individuation of our thoughts is not fully dependent on what is in our heads, but instead partly depends on the environment. Since "water" refers to different substances on earth and twin earth, "water" when used by humans on earth has different meaning than "water" when used by the beings on twin-earth. As with proper names, "water" does not have a descriptive meaning that picks out a referent. Instead, "water" directly and rigidly refers to the kind of substance dubbed, and the meaning of "water" is the substance it refers to.

## 2. MORAL SEMANTICS

G. E. Moore dissuaded philosophers from naturalistic moral realism with his Open Question Argument. Moore argued that if we identify moral goodness with some natural property, such as pleasure and the absence of pain, then it will be analytic that goodness is pleasure and the absence of pain; i.e., "goodness" and "pleasure and the absence of pain" will have the same meaning. But, he claimed, it is always an open question whether a given pleasure is good and so we cannot identify moral goodness with pleasure and the absence of pain. Since the same can be argued for any natural property that we try to identify with moral goodness, we cannot identify moral goodness with a natural property.

The developments in philosophy of language since the 1970's have made possible a new wave of moral realism. Building on the arguments of the direct designation theorists against the description theories, Boyd's semantics for moral terms, which underlies his moral

realism, rests on three main strands. First, he maintains that moral terms like “good” have a synthetic definition. With that he implies that just as “water” is not synonymous with some phrase that denotes a natural property, “good” is not synonymous with some phrase that denotes a natural property. Second, he closely follows the lead of the causal theory of reference when claiming that moral terms behave like natural kind terms. Since natural kind terms designate the same substance or properties in all possible worlds in which those substances or properties exist and so are rigid designators, moral terms are rigid designators according to Boyd.

Boyd’s third strand introduces a new wrinkle to causal theories of reference. While he holds that reference is a matter of there being certain causal connections between the use of moral terms and the relevant natural properties, Boyd’s causal theory of names differs from the standard one as developed by Kripke and Putnam. The standard view doesn’t say much about how a causal chain is transmitted, but the general view is that once the reference of a name has been fixed then it retains its reference as long as its user intends it to refer to the same object/kind as it did when she acquired the name. For example, once the reference of “water” has been fixed, “water” refers to water in all possible worlds, and whenever I use the term “water” intending the term to have its customary reference it refers to water. But Boyd’s account of reference introduces a different account of a causal connection. According to Boyd reference is essentially an epistemic notion and so the relevant causal relations constituting reference are those causal connections involved in knowledge-gathering activities. Boyd states this as follows:

“*Roughly*, and for nondegenerate cases, a term *t* refers to a kind (property, relation, etc.) *k* just in case there exist causal mechanisms whose tendency is to bring it about, over time, that what is predicated of the term *t* will be approximately true of *k* (excuse the blurring of the use-mention distinction). Such mechanisms will typically include the existence of procedures which are approximately accurate for recognizing members of instances of *k* (at least for easy cases) and which relevantly govern the use of *t*, the social transmission of certain relevantly approximately true beliefs regarding *k*, formulated as claims about *t* (again excuse the slight to the use-mention distinction), a pattern of deference to experts on *k* with respect to the use of *t*, etc. . . . When relations of this sort obtain, we may think of the properties of *k* as regulating the use of *t* (via such causal relations).”<sup>2</sup>

Horgan and Timmons formulate Boyd’s view as follows:

CRT *Causal regulation thesis*: For each moral term *t* (e.g., “good”), there is a natural property *N*, such that *N* and *N* alone causally regulates the use of *t* by humans.<sup>3</sup>

Since CRT allows one to treat moral terms as natural kind terms, they summarize the new moral semantics as follows:

*CSN Causal semantic naturalism:* Each moral term *t* rigidly designates the natural property *N* that uniquely causally regulates the use of *t* by humans.<sup>4</sup>

If CSN is true, then each moral term *t* should have a synthetically true definition whose definiens characterizes, in purely naturalistic language, the natural property that uniquely regulates the use of *t* by humans.

### 3. ENTER MORAL TWIN-EARTH

Horgan and Timmons test Boyd's moral semantics by comparing our semantic intuitions about Putnam's water/twater thought experiment and a similarly constructed thought experiment about moral terms. Recall that "water" referred to different substances on earth and twin-earth so "water" when used by humans on earth differs in meaning from "water" when used by the beings on twin-earth. In other words, our intuitive judgment when it comes to "water" is that English and Twin-English differ in meaning.

Moral twin-earth is just like earth in most respects. Twin-earthlings behave like earthlings, they make moral judgments and speak Twin-English. If earthlings were to visit twin-earth they would be strongly inclined to think that twin-earthlings use moral terms like earthlings. But there is one crucial difference between earth and twin-earth; when earthlings use moral terms, such as "good" and "right", their use of the terms is regulated by certain natural properties distinct from those that regulate the use of twin-earthlings of the same terms. Horgan and Timmons assume that earthlings' use of moral terms is regulated by some consequentialist moral properties, while twin-earthlings' use of moral terms is regulated by some deontological moral properties.

Given the similarities and differences between earth and twin-earth, what is the appropriate way to describe the difference between moral and twin-moral use of moral terms? According to Horgan and Timmons, two options are available. We could say that the differences are analogous to those between earth and twin-earth in Putnam's example, namely that moral terms used by earthlings rigidly designate the natural properties that causally regulate their use while the moral terms used by twin-earthlings rigidly designate the natural properties that causally regulate their use, so the terms refer to different properties on earth and twin-earth. If that is so, then the

moral terms used by earthlings and twin-earthlings differ in meaning. The second option is to say that moral and twin-moral terms do not differ in meaning, i.e., that they refer to the same properties, and hence that any moral disagreement that might arise between earthlings and twin-earthlings would be a genuine moral disagreement and not just disagreement in meaning.<sup>5</sup> Horgan and Timmons claim that the second option is the one that is viable, and that the first option results in the rather implausible view that inhabitants of earth and twin-earth could not engage in a genuine moral discourse about goodness. They would be talking past each other since the moral terms they used referred to different properties, i.e., had different meanings.<sup>6</sup>

Horgan and Timmons conclude from the moral twin-earth experiment that since semantic norms are tapped by human linguistic competence and since our linguistic competence is reflected in our judgment concerning twin-earth, the outcome of the thought experiment constitutes strong evidence against CSN.<sup>7</sup>

#### 4. TROUBLES FOR MORAL TWIN EARTH

Letting  $M^e$  stand for “moral term on earth”,  $M^{te}$  stand for “moral term on twin-earth”,  $P^e$  stand for “property on earth”, and  $P^{te}$  stand for “property on twin-earth”, we can state Horgan and Timmons’ Main Argument against moral semantic realism as follows.<sup>8</sup>

1. Suppose CSN is true.
2.  $M^e$  refers to  $P^e$  and  $M^{te}$  refers to  $P^{te}$ .
3.  $P^e \neq P^{te}$ .
4. So,  $M^e$  and  $M^{te}$  differ in meaning.
5. But  $M^e$  and  $M^{te}$  do not differ in meaning.
6. So, CSN is false.

Horgan and Timmons’ support for (5) rests on their claim that earthlings and twin-earthlings can have genuine moral disagreements and that they could not have genuine moral disagreements if moral and twin-moral terms differed in meaning.<sup>9</sup> Their Supporting Argument for (5) can be stated as follows.

1. If  $M^e$  and  $M^{te}$  differ in meaning, then earthlings and twin-earthlings could not have genuine moral disagreements.
2. Earthlings and twin-earthlings can have genuine moral disagreements.
3. So,  $M^e$  and  $M^{te}$  do not differ in meaning.

I will argue that the first premise in the Supporting Argument is false and that, contrary to what Horgan and Timmons claim, two parties *can* have genuine disagreements even if they are using terms that have different semantic meaning. When doing so I will employ a well-known distinction between speaker reference and semantic reference. It is the latter that Timmons and Horgan have in mind with “meaning”.<sup>10</sup>

In the philosophy of language there is a well known and accepted distinction between what a term refers to and what one uses it to refer to. What is important about the distinction for our purposes is that one can use a term to refer to an object or a property that is not the referent of that term. To use a variation of a well known example, upon seeing someone with a martini glass at a party, I might proclaim that the man drinking a martini has recently received a raise, and you might reply that Phil, who is the man with the martini glass, did not get the raise.<sup>11</sup> It turns out that Phil, the person we are talking about, does not have martini in his glass. Instead there is water in his glass. Consequently “the man drinking a martini” does not refer to Phil. The crucial point is that in spite of there being water in the glass I successfully use “the man drinking a martini” to refer to Phil. The man standing behind Phil is the one and only person at the party who is drinking a martini, and thus the person that “the man drinking a martini” refers to. Although “the man drinking a martini” does not refer to Phil, I used the phrase to refer to Phil and you picked up on my referential intention and understood who I was referring to.

Further, in order for a disagreement to arise it is not necessary that the crucial terms have a common reference or the same meaning as long as the disagreeing parties are using the terms to refer to the same thing. I might proclaim that the man drinking a martini is over 6 feet tall, and you might reply that Phil, who is the man with the martini glass, is not over 6 feet tall. We might thus get into a disagreement about Phil’s height. What enables us to disagree about Phil’s height is not that “Phil” and “the man drinking a martini” have the same reference (meaning), but rather that I successfully use “the man drinking a martini” to refer to Phil. So while “Phil” and “the man drinking a martini” refer to different objects (namely, “Phil” refers to Phil while the description refers to the man standing behind Phil who is the only man at the party drinking a martini), we can nevertheless have a disagreement about Phil while using these terms because we use them to refer to the same person and understand each other’s referential intentions.

It should now be clear that two parties can have a disagreement about something even though the terms they use in their discussion

refer to different kinds or objects, or have different semantic meaning, and that the parties can have such disagreement because speaker reference enables them to talk about objects that the terms themselves do not refer to. Further, it should also be clear that the disagreement does not have to be about the meaning of the terms and so, in Horgan and Timmons' words, the disagreement could be a genuine disagreement.

With the distinction between speaker reference and semantic reference in mind, consider whether water disagreements can arise on Putnam's twin-earth. Horgan and Timmons, presumably, would argue that if  $W^e$  and  $W^{te}$  differ in meaning, then earthlings and twin-earthlings could not have genuine water disagreements (or, more perspicuously, they could not have disagreements about the substances that  $W^e$  and  $W^{te}$  refer to).

Suppose that earthlings travel to twin-earth and there encounter twin-earthlings and twater which, remember, has all the phenomenal qualities of water with "only" its microstructure differing from water. Surely, even though "water" has different meaning on earth and twin-earth that does not have to stand in the way of the earthlings and twin-earthlings communicating about the water-like liquid on twin-earth because that will not prevent the earthlings from using "water" to refer to twater. When the earthlings, tired after the intergalactic trip, ask for water to quench their thirst, then the twin-earthlings will without hesitation bring them a glass of twater, which is what they call "water". Since twater has the same phenomenal qualities as water the earthlings will not detect that they did not get what they asked for. They might even proclaim that this is good water, using "water" to refer to what they are drinking, namely twater. Because the earthlings can use "water" to refer to twater, the earthlings and twin-earthlings can successfully communicate, or talk about the same substance when they discuss, e.g., the quality of twater, whether it would be feasible to build another twater recreational facility in town, and why twater quality is killing salmon in some river. Of course, during all these discussions all parties use the word "water" which both twin-earthlings and their visitors from earth are now using to refer to twater.

It is important that earthlings and twin-earthlings can successfully communicate about the water-like liquid, for it appears that successful communication and not shared semantic meaning of terms is a necessary condition for a disagreement to arise. The disagreement about Phil's height is a case in point. Further, it is not even a necessary condition for a disagreement to arise that the crucial terms refer at all; a point ethicists should appreciate even more in case

moral realism is false. A dramatic way to make the point is that during the 17th century scientists had serious disagreements about the nature of phlogiston and during the 1980s physicists had serious disagreements about cold fusion, and in spite of these disagreements neither “phlogiston” nor “cold fusion” have semantic meaning, since neither term refers.

Given this, it should be clear how the earthlings and twin-earthlings can disagree about “twater” even though “water” has different semantic meaning on each planet. Just as I can use “the man drinking a martini” to refer to Phil, the earthlings can use “water” to refer to twater. When they say, for example, after quenching their thirst that twin-earth has good tasting water, then they are referring to the liquid that they just drank, namely twater. And when the earthlings discuss water quality in salmon rivers on twin-earth they are referring to twater, not water. Since earthlings on twin-earth and twin-earthlings can use “water” to refer to the same kind of liquid, namely twater, they can easily have disagreements about the liquid.<sup>12</sup>

Someone might object by claiming that there is a disanalogy between the martini case and the water case. The objector might argue that when I am informed that Phil does not have a martini but water in his glass, then I will not withdraw my reference as mistaken for I intended to refer to Phil, while when an earthling is informed that the microstructure of twater is different from that of water, then she will withdraw her reference on the grounds of it being mistaken. But this objection is based on a confusion regarding semantic reference and speaker reference. A disagreement between an earthling and a twin-earthling does not have to depend upon the semantic reference of “water”, but rather upon the liquid that was being discussed and referred to, namely twater. For example, when they discussed, and argued about, why water quality was killing salmon in a river, then they were talking about and disagreeing about properties of the liquid that runs in the river, for that is what both parties referred to with “water”. Since the earthling was using “water” to refer to that liquid she could hardly proclaim that she did not intend to refer to the liquid in the river with “water”. That is, she will not withdraw her reference on the grounds of it being mistaken, for she succeeded in referring to exactly what she wanted to discuss. One way to find out what she wanted to refer to is to ask her to point at what she was referring to. She will point at twater.

If we treat moral terms like natural kind terms then we should be able to construct examples involving moral terms that are parallel to the above examples. Just as scientists disagreed about phlogiston, we should be able to have genuine disagreements about ethical issues

even if irrealism turns out to be true. And just as earthlings and twin-earthlings can have disagreements about the nature of the liquid that runs in rivers on twin-earth, earthlings on twin-earth should be able to disagree with twin-earthlings about issues that bear on twin-earthlings' well-being, i.e., about moral issues. Assuming that moral terms refer and assuming that earthlings on twin-earth and twin-earthlings can identify which properties they are talking about when using moral terms, then nothing should prevent earthlings on twin-earth from being able to use moral terms to refer to moral properties on moral twin-earth just as they can use "water" to refer to twater on twin-earth. Earthlings should then be able, via speaker reference, to use their moral terms to refer to moral properties on twin-earth.

However, if earthlings and twin-earthlings cannot identify which properties they are talking about when using moral terms, then one needs to be careful in how one constructs an example that shows that earthlings can use moral terms to refer to moral properties on twin-earth. Instead of intending to refer to a specific object or property, one would now refer to some object or property, *whatever it is*, where the reference of the name might be fixed with, e.g., a description that picks out a contingent property.<sup>13</sup> An earthling visitor could accomplish this in at least a couple of ways. First, suppose that the earthling's visit is very long. In that case it is likely that she will gradually adopt the referential intentions of the people in her new community. In that case she will gradually start to refer to moral properties on twin-earth with her moral terms.<sup>14</sup> The visitor has, in essence, acquired the twin-earthlings' moral language by gradually assimilating it and by sharing the twin-earthlings' referential intentions and so nothing stands in the way of a genuine disagreement.<sup>15</sup> Further, we can construct a second scenario where the length of the visit does not matter. Suppose that the visitor consciously decides when arriving at twin-earth that she will intend to use the moral terms she hears on twin-earth to refer to the same properties that the people on twin-earth use them to refer to. Why she decides to do so does not matter; the issue is whether it is *possible* that a genuine disagreement arise between earthlings and twin-earthlings in spite of their relevant terms differing in meaning, and not how likely it is that such a disagreement will arise. Given that she does so, she can now use her moral terms to refer to twin-earth moral properties.<sup>16</sup> Neither the earthling nor the twin-earthling might be able to identify the properties referred to, but that does not prevent them from referring to the same properties, whatever they are.<sup>17</sup> Since the

visitor and the twin-earthlings now use their moral terms to refer to the same properties, they no longer talk past each other and thus can have a genuine moral disagreement.<sup>18</sup>

Horgan and Timmons assumed that either earth and twin-earth terms differ in meaning and we cannot have a genuine disagreement about the kinds the terms refer to, or earth and twin-earth terms do not differ in meaning and we can have a genuine disagreement about the kinds the terms refer to. Their special instance of this regards moral terms. But Horgan and Timmons' argument falsely assumes that semantic meaning determines reference. It is well known in the philosophy of language that semantic meaning does not have to determine what a term is used to refer to; we have a distinction between semantic meaning and speaker meaning and I have used that distinction to show that two parties can have genuine disagreements even though the referring terms they use have different semantic meanings. Accordingly, what I have argued is that nothing prevents us from adopting a third alternative, namely that earth and twin-earth terms differ in meaning and that we nevertheless *can* have a genuine disagreement about the kinds the terms denote. I have further argued that an earthling visitor on twin-earth can use moral terms to refer to moral properties on twin-earth and so the earthling and the twin-earthlings can talk about the same properties when using moral terms.

Since two parties can have genuine disagreements while using terms that have different semantic meaning, premise (1) in Horgan and Timmons' Supporting Argument (namely, if  $M^e$  and  $M^{te}$  differ in meaning, then earthlings and twin-earthlings could not have genuine moral disagreements) is false.<sup>19</sup>

## 5. BOYD'S SEMANTICS

Since moral twin-earth fails to derail semantic moral realism the irrealist has to look for another point of attack. The most direct way to attack semantic moral realism is to question the semantic foundations on which it rests.

Horgan and Timmons formulate Boyd's basic semantic view as CSN. However, Boyd should not accept CSN as an appropriate formulation of his basic semantic view. CSN assumes that there is a single natural property,  $N$ , that causally regulates the use of each moral term by humans, while Boyd emphasizes *homeostatic clusters* of properties as regulating the use of the relevant terms.<sup>20</sup> Boyd's version has an advantage over the Horgan and Timmons version for

the following reasons; suppose that there is no *single* natural property that regulates the use of some term by humans and that instead there are several different properties that regulate the use of the term. Since there is not a single property that regulates the use of the term, the term would fail to refer on Horgan and Timmons' version, but not on Boyd's cluster formulation. CSN therefore needs to be replaced with a formulation that preserves Boyd's cluster properties. We can do that as follows:

BSN *Boyd's semantic naturalism*; A term *t* refers to *k* (kind, property, relation) just in case the properties of *k* regulate the use of *t* via a causal relation.<sup>21</sup>

BSN allows for homeostatic properties to regulate the use of a term and so it avoids the counterexample to CSN.

Boyd claims to work in the semantic tradition of Kripke and Putnam. As other causal theorists Boyd accepts rigid designators, and he has a causal element in his theory. But the similarities end there, and we should be careful to note some important dissimilarities between Boyd's semantic view and the semantic view of the other causal theorists.

Kripke and Putnam are direct designation theorists, i.e., reference is not mediated via any criteria such as a description or a set of descriptions. It is a causal chain to the baptized object that secures a name's reference. Once we have baptized an individual the name refers to that individual regardless of how she grows, develops, or changes.<sup>22</sup> In the case of substances, once the reference of the term is fixed, then the term denotes that kind regardless of the phenomenal properties the substance might have. The term *directly* designates the kind of thing that was dubbed. In the case of properties Kripke says, e.g., that heat is something we have identified as giving us a certain sensation which we call "the sensation of heat", and that we use this sensation to identify heat, subsequently fixing the reference of "heat."<sup>23</sup> Even though we used a certain sensation when we fixed the reference of "heat" we should not identify heat with whatever causes that sensation, for there are possible worlds in which we are not sensitive to heat. Again, once we have fixed the reference of the term, the reference is not mediated but is instead direct.

In contrast to Kripke and Putnam, reference for Boyd is not direct, but mediated via causal properties that regulate our use of the term. Because Boyd gives up direct designation in favor of causal regulation of our use of a term he leaves his semantic theory open to counterexamples of exactly the kind Kripke and Putnam used to argue against the description theory of kind names. I will look at a

couple of the examples that they used to show the inadequacy of the description theory and see how Boyd's view gives us the same counterintuitive results as does the description theory. The examples will demonstrate that just as a cluster of descriptions can pick out the wrong object, a cluster of properties that regulate our use of a term can pick out the wrong object. The examples will show the crucial difference between the theory of Kripke and Putnam and that of Boyd. The theory of the former has kind names track kinds in virtue of their essential properties, while Boyd's theory has names track kinds in virtue of the properties that regulate our use of the term, and those properties can be, and most often are, phenomenal and not essential properties.

Consider, as Kripke does, a counterfactual situation where fool's gold, or better yet, a substance which counterfeited the phenomenal qualities of gold but lacked its atomic structure, was actually found in all the places that now contain gold. Would we say of this counterfactual situation that gold would not have been an element, since the counterfeit substance is not an element? According to the description theory we should say that gold would not have been an element since, according to the theory, "gold" refers to the counterfeit substance as well as to gold. But Kripke disagrees, and rightly so, and thinks we should not say of the situation that gold would not have been an element.<sup>24</sup> One should not say that the substance would be gold that lacked the atomic number 79. Instead, we should say that the substance is not gold.

What should Boyd say about the same situation, given BSN? Since Boyd views kind terms as rigid designators he would have to say that the counterfeit substance is gold, since the counterfeit substance is the substance which properties causally regulate our use of the term gold. Consequently, Boyd would have us conclude that in the counterfeit situation gold is not an element. So, Kripke (and Putnam) would say, correctly, that the substance is counterfeit for gold, while Boyd and the description theorists would have to say, counter intuitively, that it is gold.<sup>25</sup>

Consider another example of Kripke's.<sup>26</sup> Suppose that we find in some until now unexplored part of the world animals that look just like tigers and have all the phenomenal properties that regulate our use of "tiger", but turn out to be reptiles when examined more closely. Do we then say that some tigers are reptiles? Since the reptile tigers satisfy most of the descriptions that constitute the meaning of "tiger", the description theory would have us say that the reptiles are tigers, while Kripke says they are not, and rightly so. Boyd would

have to say that since the properties of *both* tigers and the new-found reptiles regulate our use of “tiger”, the reptiles are tigers. Again, Boyd’s semantic theory gives us the same counterintuitive result as does the description theory.

Boyd can offer two responses to the above. First, he can say that most, but not all of the properties that regulate our use of “tiger” also belong to the reptiles and that “tiger” refers to the kind that has most of the properties that regulate our use of the term. But that will not do, since we can encounter three-legged tigers, albino tigers, or tigers that have lost their tail, in which case they lack many of the properties that regulate our use of “tiger” but still are tigers. Second, he can point out that it is an *a posteriori* question just which properties belong in the homeostatic definition of kind terms, such as “tiger”, and that once we determine which properties belong in the homeostatic definition we can settle the question of the reference of “tiger”. But this response fails to recognize that it is not the properties that belong to the homeostatic definition of a kind term that determine reference for Boyd; instead it is the properties that causally regulate our use of the term. If the phenomenal properties regulate our use of a kind term, then the phenomenal properties are the properties that are causally relevant for reference.

The problems Boyd’s semantic view encounters are very similar to the problems the description theory of names encountered at the hands of the direct designation theorists. The description theory held that a name refers to the object, or the kind of object, that best satisfied the cluster of descriptions that constituted the meaning of the term. As the direct designation theorists pointed out, this view led to names designating the “wrong” objects – e.g., counterfeit gold instead of gold and reptile tigers instead of tigers. We are seeing that Boyd’s semantic view suffers from some of the same defects as the description theory of names. Just as a cluster of descriptions can pick out the wrong object, a cluster of properties that regulate our use of a term can pick out the wrong object. The same examples that show that the description theories are semantically inadequate show that Boyd’s causal regulation view is semantically inadequate.

The first premise of Horgan and Timmons’ Main Argument relies on the truth of the CSN. I have argued that CSN should be replaced with the more plausible BSN, which has the additional virtue of having Boyd’s support. While BSN avoids some problems that CSN falls prey to, the examples that show that BSN should be rejected also show that CSN should be rejected.<sup>27</sup> The semantic foundation of Boyd’s semantic moral realism is not strong enough to support moral realism.<sup>28</sup>

## 6. BRINK'S REGULATION ACCOUNT

David Brink has expressed two worries about Boyd's semantics. First, Brink worries that if we accept a regulation theory according to which reference is determined via a causal historical chain, then

disputes about the meaning or reference of moral terms ought to be resolved not, as one would think, by moral reasoning, but by means of a historical inquiry about which features of actions, people, and institutions moral appraisers intended to pick out when those moral terms were introduced.<sup>29</sup>

This Brink finds unacceptable. Brink's second worry is that Boyd's theory gives rise to Timmons and Horgan's twin-earth arguments.<sup>30</sup>

Brink suggests that we improve Boyd's causal regulation account by introducing to it dialectic equilibrium that then plays a role in the regulation. Our ethical views reach dialectic equilibrium when we evaluate our moral principles by testing them, looking for patterns by using actual and counterfactual judgments that employ those principles. We gradually adjust our principles as coherence requires until our ethical views are in dialectical equilibrium.<sup>31</sup> Brink then gives us the counterfactual conception of regulation, according to which

a natural property N causally regulates a speaker's use of moral term "M" just in case his use of "M" would be dependent on his belief that something is N, were his beliefs in dialectical equilibrium.<sup>32</sup>

Brink's development of Boyd's theory certainly accounts for his worries about moral reasoning. But does it provide us with a reasonable *semantic* view? In order to evaluate Brink's semantic view we need to look at how it fares when dealing with the examples that gave the description theory and Boyd's theory difficulties.<sup>33</sup>

Brink's view, I believe, loosens too much the ties between names and what they name. To see this consider a variant on Kripke's gold scenario. The scenario involved a counterfactual situation where a substance which counterfeited the phenomenal qualities of gold but lacked its atomic structure was found in all the places that now contain gold. We now add to the scenario that the speakers use "gold" to refer to the counterfeited substance since they believe that the counterfeited substance is gold, and that they are in dialectic equilibrium. Would we say of this counterfactual situation that gold would not have been an element, since the counterfeit substance is not an element?

Since Brink does not specify whether or not property N is an essential property, let us first assume that N is not an essential

property and that the property in question is that of being pliable yellow metal. The speakers in the counterfactual situation certainly believe that their use of “M”, namely “gold”, is dependent on what it names being a pliable yellow metal and their beliefs are in dialectic equilibrium. Consequently, Brink’s semantic theory would have us conclude that in the counterfactual situation gold is not an element. Brink’s view thus gives us the same counterintuitive results as the description theory and Boyd’s theory.<sup>34</sup>

Now suppose that Brink specifies that the property in question has to be an essential property. That does not help much, for note that it makes no difference to the use of “gold” in the counterfactual situation whether the substance has atomic number 79. Instead what is of importance is that the speakers *believe* that it has that property. Since the speakers believe that the substance has the property, on Brink’s counterfactual conception of regulation “gold” would refer to the counterfeited substance and so the counterfeited substance would be gold and we would have to say that gold is not an element.<sup>35</sup> Further, if Brink requires that the property in question be an essential property, then he cannot account for our use of the term before we find out that the substance has that essential property, i.e., before we acquire the belief that the substance has property N. For usually we name a substance and use the name to refer to it and subsequently find out that it has some essential property. As an example, we used “water” for centuries before we discovered that water is H<sub>2</sub>O, and so we cannot account for our use of “water” before that discovery on Brink’s account.

The problem, in a nutshell, with Brink’s account is that a person’s *belief* that something is N does not make something N, even if the person’s beliefs cohere nicely and are in dialectical equilibrium. The view therefore makes the connection between names and the world too weak and has difficulties with the same kinds of examples as the description theory of names and Boyd’s regulation view.

## 7. COPP ON TWIN-EARTH

David Copp has recently replied to the twin-earth argument by suggesting that moral realists work with a semantic theory that is significantly different from the one Boyd advocates.<sup>36</sup> While Copp claims to “restrict attention to forms of synthetic semantic moral naturalism”<sup>37</sup> and thus to semantics, he argues that moral terms might be the best *translation* for the corresponding twin-moral terms,

and takes that to constitute evidence for earthlings and twin-earthlings being able to have genuine moral disagreements, for the earth moral terms and twin-earth moral terms would then have the same meaning. Copp writes: “it needs to be understood (sic) that translation is more like trying to find someone who looks enough like you to pass muster in a police lineup than it is like trying to find your identical twin.”<sup>38</sup>

The most significant wrinkle Copp introduces to the semantics of moral terms is the idea that sameness of meaning does not require sameness of reference and that meaning may be determined by functional facts rather than by causal chains. Copp considers, for example, the semantics of the term “milk” and suggests that its meaning might be provided by functional facts rather than some essential properties of milk. Copp understands his functional account as an extension of Putnam’s theory of reference of kind terms and not as a replacement of it.

Interestingly, Copp’s account of semantics flies in the face of some of Putnam’s most prominent examples, even though Copp claims to be advancing Putnamian semantics. Consider, for example, the following. Since water and twater play the same functional roles on earth and twin-earth, the best translation of “water” in Twin-English is “water” in English. Accordingly, earthlings and twin-earthlings should be able to have genuine disagreements about the liquid that fills lakes and rivers on earth and twin-earth in spite of earth having water and twin-earth water. But while it is no doubt correct that the best translation of “water” in Twin-English is “water” in English, that doesn’t change Putnam’s intuitive insight that earthlings and twin-earthlings would be talking past each other since their names refer to different substances. Finding the best translation is not sufficient for genuine disagreement, for the “best translation” might not be good enough, i.e., the best translation does not guarantee that the disagreeing parties are talking about the same thing.

Still, Copp has an interesting point when he tries to separate meaning and reference. The move is not new or novel although it hasn’t been seen before in the moral semantics debate. Neo-Fregeans have for some time accepted a causal theory of reference where names nevertheless have meaning, or sense, that does not determine reference. Their motivation for this divergence of sense and reference is to account for information value of, e.g., true identity statements with different but codesignative names. Neo-Fregeans tend to accept a causal theory of reference, which is the key to their semantic

account, or how the language connects with the world. While they use sense to account for information value, the sense does not determine reference. Copp's twist is to advocate *functional roles* instead of a causal chain as providing semantic meaning (and thus the reference of the term). But functional roles face the same problems as descriptive contents when it comes to providing reference. Suppose, for example, that we decide that oil is whatever it is that provides us with fuel for our cars and lubrication for our motors, i.e., suppose we define "oil" functionally. Suppose also that in the not so distant future all fuel and lubrication, i.e., all oil, will come from soy beans. It seems clear that we and our descendants are referring to different substances with "oil" and hence that in a significant way we are not talking about the same thing when using the term. Incidentally, it seems that Putnam would agree, for he writes: "If there is a hidden structure, then generally it determines what it is to be a member of the natural kind, not only in the actual world, but in all possible worlds."<sup>39</sup> Different substances and different properties can play the same functional roles, and if we want to talk about the substances and properties *themselves* instead of whatever it might be that plays their functional role, then Copp's semantic twist fails to do the job. In other words, Copp has not provided us with an account of *semantic meaning*, or reference, that avoids problems of the type Kripke and Putnam generated for the description theories.

#### 8. THE PROSPECT FOR SEMANTIC MORAL REALISM

Boyd's, Brink's, and Copp's attempts to construe semantic moral realism are arguably the best attempts at doing so at the present time. Their project is not to use semantics to argue that moral realism is true. Instead they are trying to provide a package comprising metaphysics, semantics, and epistemology that is overall coherent and plausible. A key element in the total package is the semantic view, for there is a serious sense in which the semantic account provided is the foundation upon which semantic moral realism has to rest. If the semantic account given does not stand up to scrutiny, then *semantic moral realism* is in trouble, for then the overall package lacks plausibility. At the core of the semantic views of the semantic moral realists is the claim that moral terms rigidly designate natural properties. But all three attempts we have looked at to provide the details of an account that supports such a claim fail. Since Boyd's, Brink's, and Copp's attempts all have faulty semantic foundation, and these

are arguably the best attempts so far to provide a semantic account for semantic moral realism, one has to wonder whether semantic moral realism has any future.

In order to have viable realist moral semantics we need to model it on the success of the direct designation theory, which is arguably the best semantic theory of names we have at this time. The direct designation theory assumes that in order to fix the reference of a kind term one needs to establish the existence of the property or kind one is naming. For example, we would not want to admit that phlogiston was a referring term, or that at one time it referred and now it doesn't do so. We do not want to say that 17th century scientists fixed the reference of "phlogiston" which, by the way, fails to refer. When one fixes reference one has to fix it onto something; i.e., *first* one establishes the existence of the kind and *then* one names it, or *first* one establishes the existence of the relevant properties and *then* one names them. But if one needs to establish the existence of an object or a property named before naming it, then one needs to establish the existence of moral properties prior to fixing the reference of moral terms. Before one can even assume that the direct designation theory, or a causal theory of names, gives us the appropriate semantic framework for moral semantics, one needs to establish the existence of the entities to be named.

It is widely acknowledged that the direct designation theory, and any causal theory of reference for that matter, faces severe difficulties with non-referring names, and so one cannot assume that such a semantic theory provides the correct framework when one does not know about the ontological status of the objects to be named. Given the difficulties that the direct designation theory has with empty terms, we cannot assume that the direct designation theory provides us with the appropriate semantics for moral terms when we don't know whether or not they are referring terms; that is, when we don't know whether or not the properties presumably referred to exist. A directly designating term has to refer to *something*, even though we might not know the nature of that something and have to find out about it later. But, since moral realism might be wrong, it is not clear that moral terms refer to anything at all. In fact, "good" and other moral terms might be more on par with "phlogiston" than "water".

The pioneers of the causal theory of reference gave examples where people introduced terms for objects they were not acquainted with. But even then, they knew that *something* was at the end of the causal chain. Donnellan, for example, suggests that Leverrier introduced "Neptune" as follows: "If the planet which caused such and

such discrepancies in the orbit of Uranus exists, then Neptune is the planet which caused such and such discrepancies on the orbit of Uranus.” But even Donnellan suggests in the following discussion that we might not have *named* a planet; instead we might only have *reserved* a name for it. This is compatible with the view that we talk *as if* the planet exists and we talk *as if* we have named it. Similarly, the scientists who theorized about phlogiston certainly talked *as if* it existed, and in spite of that it would be wrong to claim that they named phlogiston, for it doesn’t exist. How are we to explain this? It is possible that Azzouni is correct when he argues that the solution that suggests itself is that we *posit* the existence of the (theoretical) objects we talk about and that we subsequently take our terms to so refer.<sup>40</sup> Similarly, Mackie suggested that we talk *as if* moral properties exist though, in fact, they don’t. But this only shows that we can use what *seems* to be the language of realism when, in fact, we are positing existence and the objects or properties being talked about may not, and sometimes do not exist. As Azzouni points out, our causal reach is far less than our referential bluff.

Perhaps, then, we can use the language of moral realism within the framework of direct designation when discussing moral realism. But the views of Boyd, Copp, and Brink seem to suggest otherwise. With their focus on homeostatic properties and functional roles they seem to indicate that moral properties are not of the kind that can be easily identified, as water, e.g., can be identified as H<sub>2</sub>O. Instead the suggestion is that we are dealing with complex properties, on par with the concept of, e.g., health, where no single combination is necessary and/or jointly sufficient for something being healthy and, similarly, no single combination of properties might be necessary and/or jointly sufficient for, e.g., something to be good. Or we might be dealing with functional properties where, e.g., no single substance might have the only claim to be milk. While the direct designation theory has a number of strengths, it does have difficulties dealing with the kind of fluid complexes that moral properties seem to be. Boyd, of course, realized this and resorted to devising his own semantic account designed to deal with homeostatic properties.

If the direct designation theory does not provide the right semantics for moral terms, then we can try to look to the description theory. But the description theory fares no better than does the direct designation theory. Boyd is quite correct when he points out that we do not treat moral terms as having analytic definitions. Instead it is very reasonable, as Boyd and Brink point

out, that we gradually discover and refine the semantic meaning of moral terms, as we can gradually discover and refine the semantic meaning of other natural kind terms as we uncover the nature of the kind named, and that we adjust their meanings as we come to know more. But if neither theory provides us with an adequate account of the semantics of moral terms, then how should we understand our use of the terms?

The way we use theoretical terms, including terms for moral properties, assuming those properties exist, seems to fit neither the direct designation theory nor the description theory. Instead, our use of the terms seems to go, roughly, as follows. We seem to have an idea of something we are talking about and, perhaps, searching for. We might use a name when we talk about the supposed entity, like “phlogiston” or “quark” or “just”, but we are not sure that the terms refer to anything or, in case they do, we are not sure what exactly they refer to. Still, we may have an idea of the properties the supposed entity has and that idea gives us a “loose” meaning for the term. As the search continues and we gradually hone in on the entity we realize that we were wrong about some of the properties we thought it had, and we subsequently refine our ideas of the supposed entity and the semantic meaning of the term we use when discussing it. Eventually we might realize that the entity we were searching for and thought we were talking about doesn’t exist, as happened with phlogiston, or we might successfully hone in on the object, as happened with quarks. In the former case we have a non-referring term, and in the latter a referring term which semantic meaning has gradually been clarified via empirical research. But the nature of moral properties seems to be significantly different from, e.g., the nature of a quark, the former probably having homeostatic qualities. It is quite likely, as I have pointed out, that due to that nature we might not be able to find a causal underpinning for reference of those terms. As a matter of fact, if moral twin-earth represents a plausible scenario, then that alone shows that we can have radically different causal underpinnings for otherwise identical moral language and moral judgments. The account suggested here has therefore strong unrealistic and non-naturalistic elements, as it suggests that while our use of moral terms might be guided by epistemic considerations, we might not be able to forge a causal referential link to moral properties.

At the present time we are still not sure whether or not moral terms have semantic meaning. The moral realists are still searching and trying to hone in on natural moral properties, but they might

have to admit, as the early chemists who searched for phlogiston, that they were led on a wild goose chase and that irrealism is true. Regardless of whether or not we find moral properties we can continue to use moral terms as we do, gradually refining the meaning of the terms as we see fit. If it turns out that irrealism is true, then, from a semantic point of view, not much is lost.

## 9. CONCLUSION

The argument I have offered against the new moral realism avoids the twin-earth arguments and twin-earth intuitions that face problems of their own. The argument also demonstrates that there are significant differences between the standard causal theory of reference and Boyd's version of the causal theory, and that the differences are such that Boyd faces problems that are similar to the problems that defeated the description theory of reference. Further, the semantic views Brink and Copp present fare no better than Boyd's theory. The most plausible way to account for our use of moral terms is to give up the view that moral terms directly designate moral properties. But once the semantic realist loses direct reference from her arsenal, and with that synthetic definitions, Moore's open question argument remains a serious challenge to naturalistic ethical realism.

## ACKNOWLEDGEMENTS

Thanks go to Margaret Holmgren and Peter Vranas for helpful comments and discussions. I also thank Pat Francken, Rod Bertolet, and referees for *Erkenntnis* for their comments.

## NOTES

- <sup>1</sup> "Troubles for New Wave Moral Semantics: The 'Open Question Argument' Revived," *Philosophical Papers*, 1992 (153–175), "Troubles on Moral Twin Earth: Moral Queerness Revived," *Synthese*, 1992 (221–260), "New Wave Moral Realism Meets Moral Twin Earth," *Journal of Philosophical Research*, 1990–1991, (447–472). See also Timmons' *Morality Without Foundations: A Defense of Ethical Contextualism*, (New York: Oxford University Press): 1999, and "On The Epistemic Status of Considered Moral Judgments," *The Southern Journal of Philosophy*, 1990 (supplement).

- <sup>2</sup> Richard Boyd, "How To Be a Moral Realist," *Essays on Moral Realism*, ed. Geoffrey Sayre-McCord, (Ithaca: Cornell University Press, 1988): 181–228, p. 195.
- <sup>3</sup> Horgan and Timmons, "Troubles for New Wave Moral Semantics," p. 159, Timmons, *Morality Without Foundations: A Defense of Ethical Contextualism*, p. 58.
- <sup>4</sup> Horgan and Timmons, "Troubles for New Wave Moral Semantics," p. 159, Timmons, *Morality Without Foundations: A Defense of Ethical Contextualism*, p. 58.
- <sup>5</sup> Horgan and Timmons, "Troubles for New Wave Moral Semantics," p. 165.
- <sup>6</sup> Horgan and Timmons, "Troubles for New Wave Moral Semantics," p. 166.
- <sup>7</sup> Horgan and Timmons, "Troubles for New Wave Moral Semantics," p. 166.
- <sup>8</sup> We should be careful to note that "meaning", in this discussion, is to be read as 'semantic meaning'. That is, in the spirit of direct designation, which Boyd builds on, the meaning of a term is what it denotes or refers to. Also note that while a term has a reference, it may be *used* to refer to something else. But more on that later when I discuss the relevance of speaker reference and semantic reference.
- <sup>9</sup> Here I am relying on the argument in "Troubles for New Wave Moral Semantics." In other places Horgan and Timmons rely on an intuition that tells us that the disagreement between earthlings and twin-earthlings is about their *beliefs* and ultimately their *moral theories* (see e.g., "Troubles on Moral Twin Earth," pp. 247–248). I do not find that setup very convincing. Assuming that Boyd's semantics is correct and thus that an earthling on twin-earth refers to the same moral properties as the twin-earthling, our intuition tells us that there is nevertheless a moral disagreement between the two. But, assuming that the earthling has adopted a consequentialist theory and the twin-earthling a deontological theory, one would assume that the theoretical framework resists any quick changes even though the moral terms now refer to different properties than before. But while the theoretical framework would resist any quick changes, it should, as time goes by, adjust to the new reality. Eric Kraemer mentions problems with moral intuitions in "On the Moral Twin Earth Challenge to New-Wave Moral Realism," *Journal of Philosophical Research* 1990–1991, pp. 467–472.
- <sup>10</sup> Horgan and Timmons have another line of argument where they claim that when presented with the twin-earth thought experiment then our "intuition packet" kicks in, much as it does with Putnam's twin-earth, and informs us that earthlings and twin-earthlings are talking about the same things and can have genuine disagreements. I discuss this line of reasoning, which I do not find very convincing, in "Moral Twin-Earth: The Intuitive Argument," *Southwest Philosophy Review*, Vol. 19 (2003): 115–124.
- <sup>11</sup> The example is adapted from K. Donnellan's "Reference and Definite Descriptions," *Philosophical Review*, 1966.
- <sup>12</sup> In Boyd's discussion, as well as in Horgan and Timmon's discussion, we do not know the precise nature of the moral properties that underlie the causal regulation of the use of our moral terms. That is, we assume that earthlings use is regulated by *some* consequentialist moral properties while the twin-earthlings use is regulated by *some* deontological moral properties and at the same time we do not know the exact nature of these properties. One of the projects for the moral realist is to uncover those properties and, of course, it is a strength of synthetic definitions that we can conduct investigations to uncover the precise meaning of the terms being

defined. In order to keep the analogy in our water/twater talk, we should assume that while the substances are different, the discussants do not know the exact nature of the properties that causally regulate their use of “water”.

<sup>13</sup> Compare, e.g., how Donnellan and Kripke treat the reference fixing of ‘Neptune’, where the reference of the name was fixed with “If the planet which caused such and such discrepancies in the orbit of Uranus exists, then Neptune is the planet which caused such and such discrepancies on the orbit of Uranus.”

<sup>14</sup> See Tyler Burge on fast switching and slow switching when discussing intergalactic travels. During fast switching the traveler stops for a very short time and might not even realize that she is in a different place. During slow switching the traveler stays for a long time; long enough for her thoughts to be about the kinds/objects at her destination. See Burge’s “Individualism and Self-Knowledge,” *Journal of Philosophy*, 85/11 (1988), 649–663. It appears to me that if the traveler stays long enough for her thoughts to be about the kinds on twin-earth, then her terms also refer to those kinds.

<sup>15</sup> Brink fails to notice the significant difference between the slow switching case and the fast switching case and assumes that as soon as the earthling comes to twin earth she shares their referential intentions. See Brink, pp. 174–175 in David O. Brink “Realism, Naturalism, and Moral Semantics,” *Social Philosophy and Policy*, vol. 18, 2001: pp. 154–176.

<sup>16</sup> Suppose Horgan and Timmons try to counter as follows: “Suppose we are trying to decide whether to find an innocent person guilty and by so doing preventing riots or finding him innocent and by doing so risk the lives of others due to riots. We are disagreeing over what the right thing to do is, and that means that we are disagreeing about what the word ‘right’ refers to. But since the speakers have no independent access to what they are referring to, as in the Phil case, we cannot rely on speaker reference to ground substantial moral disagreements.”

There are two problems with this reply. First, as I argue below, the speakers can refer to the same thing even though they can not identify that thing. Second, I do not think we can assume that a disagreement about what to do means that we disagree about what ‘right’ refers to. For example, there are many variations of consequentialist theories and they do not always yield the same results in the same moral situation, and in spite of that they are all, presumably, based on consequentialist properties. The point is that even when they are known, moral properties by themselves, do not dictate the outcome of a moral situation. Theoretical issues might play a decisive role in the ultimate outcome.

<sup>17</sup> Compare this with, e.g., Donnellan’s Neptune example. There Leverrier, supposedly, can use ‘Neptune’ to refer to Neptune in spite of not being able to identify Neptune. Suppose Leverrier passes the name on to his assistant. Then both can use ‘Neptune’ to refer to the same object and still neither can identify the object referred to.

<sup>18</sup> Does this mean that ‘water’ now has the same meaning in the mouth of the visitor and the twin-earthlings? Not necessarily. The semantic meaning of ‘water’ in the mouth of the earthling might still be water while she uses the term on twin-earth to refer to twater. The distinction between speaker reference and semantic reference can be extended to names. Consider a scenario where I see Peter working in a field and you mistake him for John. You say “John is working hard today”. I agree that the person in the field is working hard, and not bothering correcting you, say “John is almost done bailing”. Even though I use ‘John’ to refer to Peter,

the meaning of 'John' is still John. The referential intentions have the semantic meaning and reference come apart.

- <sup>19</sup> M. Holmgren, in "The Poverty of Naturalistic Moral Realism: Comments on Timmons", *The Southern Journal of Philosophy*, 1990, pp. 131–135, argues that the difference between moral terms and natural kind terms is that the former are evaluative and the latter not and so the identification of a natural moral property would not give us reasons to perform one action rather than another. E. H. Campel, in "Ethics, Reference, and Natural Kinds," *Philosophical Papers*, 1997, pp. 147–163 makes the point that we don't seem to intend to refer to naturalistic moral properties with out use of moral terms. Both points seem to beg the question against the moral realist. Against Holmgren, the realist might point out that moral terms can be evaluative and naturalistic and that when we truly say that one act is better than another act then the better act has more, or stronger natural moral properties than the worse act, where one act could have more, or stronger natural moral properties than another if it has more properties from the relevant homeostatic cluster, or more central properties from the relevant homeostatic cluster than the other. And against Campel the realist could point out that the even though we don't specifically intend to refer to natural moral properties, the terms we use nevertheless denote those properties, and that in the absence of a special intention we use terms to refer to their customary reference, so those are the properties we are talking about.
- <sup>20</sup> Boyd uses the term 'healthy' as an illustration of homeostatic cluster definition, although he does not discuss whether or not 'health' is a full-blown kind term.
- <sup>21</sup> Boyd, p. 195.
- <sup>22</sup> For reference of kind names, see H. Putnam, "It Ain't Necessarily So," *Journal of Philosophy*, 1962, pp. 658–671 and "The Meaning of 'Meaning'," *Language, Mind and Knowledge*, ed. K. Gunderson, (Minneapolis: University of Minnesota Press, 1975), S. Kripke, *Naming and Necessity*, (Cambridge: Harvard University Press, 1980), p. 122.
- <sup>23</sup> Kripke, p. 131.
- <sup>24</sup> Kripke, p. 124.
- <sup>25</sup> Suppose Boyd replies to the argument by simply conceding that in the counterfactual situation 'gold' refers to fools gold, since it is gold that actually regulates our use of 'gold'. The reply would fail on two grounds: first, it would give up the view that moral terms are rigid designators, and thus he would no longer have synthetic definitions of moral terms. Second, it is true that gold *actually* regulates our use of gold, but the counterfactual world is not the actual world. Hence the difference.
- <sup>26</sup> Kripke, p. 120. The example is slightly modified to assure that we established the use of 'tiger' before we found the reptile tigers.
- <sup>27</sup> Instead of finding a cluster of properties that regulate our use of a term, we would need to find a single property that does so. We then set up a counterfactual situation where a different substance has that property (as, e.g., the counterfactual gold). The problem arises in the same way as before.
- <sup>28</sup> At this point one might ask why Boyd doesn't simply revert to the direct designation theory as advanced by, e.g., Kripke and Putnam and without the causal regulation element that he builds into it. The changes in Boyd's account would be too sweeping for him to take that rout. For example, he would have to give up the homeostatic conception of moral terms, as they play a key role in his semantic account of moral terms. Little would remain of his original account.

- <sup>29</sup> Brink, p. 164.
- <sup>30</sup> Brink, pp. 165–166.
- <sup>31</sup> Brink, pp. 168–169.
- <sup>32</sup> Brink, p. 169.
- <sup>33</sup> Geoffrey Sayre-McCord provides an important precursor to Brink's view in "'Good' on Twin-Earth", *Philosophical Issues*, vol. 8 (1997), pp. 267–292. In that paper Sayre-McCord retains causal regulation, as does Brink. He also emphasizes the role normative theory should play in an account of the reference of moral terms and so he seems to emphasize equilibrium and its role in determining the reference of moral terms. It seems that Sayre-McCord's view, with its causal regulation and equilibrium, faces the same problems as does Brink's view; namely, the connection between names and the world is weakened too much and, as a result, the view falls prey to Kripke/Putnam like counterexamples.
- <sup>34</sup> There is a very close parallel between the arguments here and the arguments Kripke, Putnam, and others gave against the description theory.
- <sup>35</sup> It is important to recognize that coherence does not guarantee truth and so the fact that one's beliefs cohere nicely is compatible with some of them, even most of them, being false.
- <sup>36</sup> Copp, David, "Milk, Honey and the Good Life on Moral Twin Earth," *Synthese* 124 (2000): 113–137. Horgan and Timmons reply to Copp in "Copping Out on Moral Twin Earth," *Synthese* 124 (2000): 139–152.
- <sup>37</sup> Copp, p. 120.
- <sup>38</sup> Copp, pp. 121–122.
- <sup>39</sup> Putnam, Hilary, "The Meaning of 'Meaning'," *Mind, Language and Reality: Philosophical Papers* vol. 2, (Cambridge: Cambridge University Press, 1975): 215–271, p. 241.
- <sup>40</sup> Azzouni, Jody, *Knowledge and Reference in Empirical Science*, (New York: Routledge, 2000).

Department of Philosophy  
 Iowa State University  
 Catt Hall  
 Ames IA 50011  
 United States  
 E-mail: geirsson@iastate.edu

Manuscript submitted 13 November 2003

Final version received 9 August 2004