

The Revenge of Moral Twin Earth

Abstract

In this paper I revisit an important response to the Moral Twin Earth (MTE) challenge: The Common Functional Role strategy (CFR). I argue that CFR is open to a revenge problem. MTE-cases allegedly show that two linguistic communities can be in genuine disagreement even when they are regulated by distinct families of properties. CFR provides a way to reconcile the intuition that the two communities are in genuine disagreement with the claim that the use of moral terms by both communities is causally regulated by different families of properties. This is done by identifying a functional role that those families of properties both fulfil. Still, even if CFR is successful, its proponents need to face a serious revenge problem. Roughly, it could be that the families of properties that regulate each community are equally perfect realizers of the relevant higher-order functional state. I suggest that the proponent of CFR faces a dilemma: either CFR has controversial implications about first-order moral theory, or CFR needs to be coupled with substantive and parochial empirical/metaphysical assumptions to avoid those implications.

Keywords: Moral Twin Earth, Common Functional Role, Moral Functionalism, Moral Disagreement, Moral Realism

1. Moral Twin Earth and Causal Metasemantics

Certain versions of moral realism appeal to causal metasemantics to make sense of moral discourse. The Moral Twin Earth (MTE) thought experiment is a powerful challenge to those accounts. The most influential account, and the main target of MTE cases, can be found in Boyd's work (1988; 2001a; 2001b).¹ Roughly, the idea is this. True moral beliefs are true in virtue of the fact that the terms employed by those beliefs are causally regulated by collections of stance-independent moral properties. Horgan & Timmons, the originators of the MTE challenge, present Boyd's causal metasemantics in the following way (H&T 2013: 2):

(CM) For every moral term *t* (e.g. “wrong”, etc.), there is a collection of natural properties *N*, such that *N*, and *N* alone, causally regulates *t*'s use.

H&T argue that CM leads to counterintuitive results. To properly set the stage, consider Putnam's original twin Earth scenario (1975). When Earthlings use the term “water”, their use is regulated by instances of

¹ Causal metasemantics are typically used by naturalists. This explains why, historically, MTE cases have naturalist versions of moral realism as their target. But such metasemantics can also be used by anti-naturalist moral realists. For example, Adams's (1999) supernaturalist moral realism has a structurally identical metasemantics to Boyd's.

H₂O. On the other hand, by stipulation, when Twin-Earthlings (henceforth, T-Earthlings) use “water”, their use is regulated by instances of XYZ. So, when Earthlings and T-Earthlings disagree about the use of “water”, their disagreement is illusory given that each community refers to a different entity. Still, this result seems unproblematic. Our intuitions agree that the relevant disagreement is not genuine.

Moral naturalists like Boyd claim that CM applies to moral terms in the same way it applies to terms like “water”. So non-genuine disagreement about a moral term like “wrong” should be analogous to non-genuine disagreement about terms like “water”, as presented by Putnam’s original Twin-Earth case. H&T argue that this is not the case. To illustrate this, they present the following scenario: When Earthlings use “wrong”, their use is regulated by properties that satisfy a consequentialist theory (e.g. properties whose instantiation decreases overall utility). On the other hand, by stipulation, when T-Earthlings use “wrong”, their use is regulated by properties that satisfy a deontological theory (i.e. properties whose instantiation violate some deontic constraint like the categorical imperative).²

Now imagine an Earthling and a T-Earthling having a moral disagreement (perhaps, about a trolley case). According to our intuitions, and textbook moral philosophy, such a disagreement is genuine. But it seems that CM cannot accommodate this result. Recall that, by stipulation, Earthlings and T-Earthlings are causally regulated by different collections of properties.³ Assuming that sameness of reference is necessary for genuine disagreement, it follows that Earthlings and T-Earthlings are not in genuine disagreement.⁴ The two parties are merely talking past each other.

Typical responses to H&T’s challenge include denying the relevant intuitions by biting the bullet or arguing that MTE cases are incoherent (H&T 2003: 5-6). In this paper I will focus on a different kind of response: The Common Functional Role strategy. I will argue that it faces a revenge problem, so far unaddressed in the relevant literature.

2. The Common Functional Role strategy

According to the Common Functional Role strategy (CFR) there is a way to reconcile CM with the possibility of genuine moral disagreement between Earthlings and T-Earthlings.⁵ How could this be given

² Someone might be hesitant in combining deontology with Boyd’s moral realism (although see sec. 4). Still, note that we can construct an equally powerful MTE case involving different moral theories (e.g. Freiman [2014] appeals to virtue ethics).

³ For the rest of this paper, I will use the phrase “Earthlings are regulated by some properties” as a shorthand for saying that the *use of their terms* is regulated by these properties.

⁴ For dialectical reasons, I will grant this controversial assumption (cf. Plunkett & Sandell 2013). At any rate, the same argument could be presented (by adding further epicycles) by appealing to intuitions concerning whether the two communities track the same *referent* (see Väyrynen 2018).

⁵ The CFR response has been proposed, in different ways, by Kraemer (1991: 469), Lawrence et al. (1999: 158), Copp (1995; 2000: 133; 2007), Brink (2001: 172-3), Levy (2010: 142, 142 fn. 1). It is also recognized by its critics (e.g. Rubin

that, by stipulation, Earthlings and T-Earthlings are regulated by different collections of properties? The idea is this: The kinds of properties that regulate Earthlings and T-Earthlings are distinct but exhibit functional resemblance. In more precise terms, consequentialist properties and deontological properties *functionally realize* the same higher-order property.⁶ In this sense, both Earthlings and T-Earthlings are causally regulated by the same, higher-order, property.⁷

What property might that be? Plausible candidates include the contribution to some notion of well-being (Brink 1984), societal stability (Copp 2007), or the satisfaction of important human needs (Boyd 1988: 329). For illustration purposes I will focus on need-satisfaction. The resulting picture seems to be this: Earthlings are regulated by consequentialist properties (henceforth, CP) and T-Earthlings are regulated by deontological properties (henceforth, DP). Still, sameness of reference is preserved given that both CP and DP are genuine realizers of human needs. This way, one can respect the conditions of H&T's thought experiment and still claim that Earthlings and T-Earthlings are in genuine disagreement given that they are regulated by the same second-order property (i.e. need-satisfaction; Needs henceforth).

To illustrate the idea that a set of properties can functionally realize Needs, consider a non-moral type of need. Let's suppose that in the Americas, prior to the Age of Exploration, *nutritional* needs were realized by the consumption of a combination of things like potatoes, corn, quinoa, tomatoes, and game meat, while in what is today Southeast Asia they were realized by the consumption a combination of things like rice, fish, greens, roots (e.g., lotus root), melons, and fungi.⁸ In such a case, the satisfaction of nutritional needs can be functionally realized by two distinct sets of first-order properties. In other words, there are two distinct types of diets which, if adhered to, would meet our nutritional needs.

Now consider Needs. A more concrete *moral* example involving need-satisfaction might be a scenario where Earthlings follow a broadly rule-utilitarian moral code, whereas T-Earthlings adhere to Rossian *prima facie* duties. The former society would follow rules that, adhering to them, tends to maximize utility, whereas the latter is based on a set of foundational duties, such as fidelity, reparation, and beneficence. These two societies would be substantially different since, by stipulation, the respective

2008: sec. 7.3.). Traces of CFR can also be found in Merli (2002: 236) and Sonderholm (2013: 86). It is also a natural response for moral functionalists who adopt CM (e.g. Brink 1984: 122-3; Railton 1993).

⁶ The exact metaphysics of the realization relation are not relevant at this point. For an overview of the different options see Baysan (2015).

⁷ It could be argued that it is unclear how a *realized*, higher-order, property can causally regulate one's use of moral terms. In response, all that is required is that such properties are causally efficacious. There are at least two ways this can happen. First, such realized properties can be identical to the family of properties that realize them (as per, so-called, *filler* functionalism). Secondly, such realized properties can be ontologically distinct from their realizers but can nonetheless *inherit* their causal profile. In this sense, I disagree with Rubin's (2008: 325 fn. 27) claim that proponents of CFR ignore H&T's stipulation that different first-order, non-functional, sets properties causally regulate the relevant terms used by the two communities. That stipulation is compatible with the picture that CFR suggests.

⁸ [Acknowledgements].

communities follow distinct moral codes. But, according to CFR, this simply illustrates that there two distinct, but legitimate, ways of bringing about the satisfaction of Needs.⁹

According to this picture there is just something about the nature of our Needs that they can be satisfied in multiple ways. The Boydian idea appealed to here is that there is a fixed, mutually supporting, cluster of important human needs such as “the need for love and friendship” or the need “to engage in cooperative efforts” (Boyd 1988: 329).¹⁰ To quote Boyd: “Actions, policies, character traits, etc. are morally good to the extent to which they tend to foster the realization of these goods” (1988: 329).

Take friendship for example: it could be that there is something about its nature which allows for many type-distinct paradigms of friendship. Perhaps, there are types of friendship where the promise-keeping duty should be prioritized even when it leads to unpleasant consequences. And perhaps this is so even though that same need for friendship could be met by a paradigm which is more tolerant towards the violation of promises (perhaps, in the presence of certain outweighing reasons involving the relief of a certain level of distress).

Of course, things might not be as such in the sense that there is one, and only one, type of friendship (and, extrapolating, one type of moral code which brings about the satisfaction of Needs) (more on this later). Compare with nutritional needs: it could have been the case that our nutrition needs are as such that we should follow one, and only one, type of diet.

Still, I am not particularly interested in defending CFR here. What I am interested in is the revenge problem I claim it faces. For my purposes it suffices that CFR is an important strategy against MTE cases and it has a sufficiently high degree of plausibility. Also, it seems to be a straight-forward way for proponents of CM to deal with MTE cases without denying the relevant intuitions: CFR, if successful, manages to establish a common referential target for both communities. Earthlings are regulated by CP and T-Earthlings are regulated by DP. The fact that both CP and DP functionally realize Needs makes sure that both parties are talking about the same thing (even though they diverge at the first-order level).¹¹

I take the Boydian picture sketched above to be quite plausible, as well as the idea that there could be *multiple* ways of bringing about the satisfaction of Needs.¹² But even if one is unconvinced by the appeal

⁹ Whether deontology is consequentializable is irrelevant at this point. Instead of CP and DP one could simply appeal to two extensionally divergent first-order moral theories.

¹⁰ For Boyd, figuring out the exact details of these needs is an empirical issue (Boyd 1988: 329). Boyd also suggests that the relevant idea of need-satisfaction applies to *communities* rather than particular individuals (1988: 353). This allows for a picture where there are particular trade-offs in the satisfaction of the needs of individuals, even though the needs of the community *as a whole* are satisfied.

¹¹ So-called ‘Deniers’ (H&T 2013: 5) include Dowell (2016) and Wisdom (2021).

¹² To clarify, the scenario being presented here is not one according to which Needs is *actually* perfectly realized by CP and/or DP. This is because, plausibly, neither CP or DP (or whichever properties are posited by the extensionally divergent first-order theories of one’s choice) are *actually* instantiated. It is perfectly coherent for a community to be causally regulated by CP (qua perfect realizer of Needs) even if the full array of properties figuring in CP is not instantiated.

to Needs, this has no *particular* bearing on the CFR strategy. Perhaps, as mentioned, the common functional role should be construed in terms of another higher-order phenomenon such as the “problem of sociality”, roughly, the problem of finding ways of avoiding conflict in a societal context: again, perhaps *both* a rule-utilitarian moral code and a moral code based on Rossian prima facie duties would solve the problem of sociality, albeit in different ways (by instantiating different types of societies). Perhaps there are reasons why the picture is better than the Boydian one. Still, in this paper I will focus on CFR as a general type of strategy, while framing things in terms of Needs for illustration purposes.

3. Revenge

There are at least two ways to cash out the claim that CP and DP are both realizers of Needs. The first way concerns a scenario that involves referential *convergence* between the two communities. For example, it could be that, even though CP and DP are realizers of Needs, one of them is an *imperfect* realizer of Needs. Let a perfect realizer be a family of properties whose every member sufficiently contributes towards the realization of the relevant higher-order phenomenon. An imperfect realizer, on the other hand, is a family of properties that, broadly speaking, brings about the relevant phenomenon in an imperfect manner.¹³

There are potentially many ways in which an imperfect realization can occur. Still, I will focus on the one that is relevant to my argument. Consider a family of properties that *lacks* certain properties that are essential for the implementation of the relevant higher-order phenomenon.¹⁴ A potential non-moral example of an imperfect realizer could involve the realization of a pain-state without the generation of its accompanying phenomenal experience. This seems to be what is behind cases involving philosophical zombies (Chalmers 1996). In such cases, the relevant physical states actively contribute towards the realization of their corresponding mental states. It is just that there is something missing that would otherwise generate the relevant phenomenal state.¹⁵ So, it could be that CP are imperfect realizers of Needs whereas DP realize Needs perfectly (or, perhaps, they are both imperfect realizers to different degrees). In

¹³ The perfect-imperfect distinction does not directly map onto the standard distinction between a *total* and a *core* realizer (e.g. Shoemaker 2007). Total realizers are sufficient for what they produce. But notice that this leaves open the possibility of a total realizer with superfluous properties (see note 12). Similarly, a core realizer is, roughly, the most salient and central part of a total realizer. In this sense, a core realizer is a particular instance of an imperfect realizer.

¹⁴ For this reason, my use of the perfect-imperfect distinction is somewhat different from Nola & Braddon-Mitchell’s version (2009: 8-9). Their version identifies perfect and imperfect realizers with natural and non-natural properties (as per Lewis [1983]), respectively. Not-so-natural properties are, roughly, gerrymandered and heterogeneous. But, in my sense, a realizer can be imperfect but not gerrymandered (e.g. a realizer might be homogenous while still having some missing properties which would otherwise fully realize the relevant higher-order phenomenon in its most complete sense).

¹⁵ Another case might be the following: There could be a family of properties that *would* be sufficient for an accurate realization of some higher-order phenomenon, if it weren’t for some *superfluous* properties. Again, consider the pain case. There could be a scenario of pain-realization where the relevant phenomenal experience of pain is accompanied by a phenomenal experience of tasting liquorice.

this case, if ideal epistemic conditions apply, both communities will converge towards the perfect realizer of Needs. Still, even without perfect convergence, the two communities can disagree meaningfully.¹⁶

This first way of understanding the claim that CP and DP functionally realize Needs is presumably what proponents of CFR have in mind when they motivate their view. Another possibility, however, involves convergence at the functional level but divergence at the first-order level. Consider MTE again. Both communities are regulated by the same functional property (the satisfaction of Needs). However, this time, the realizer types are both perfect *and* distinct from one another. In this scenario, the first-order divergence between the two communities will remain *even under* ideal epistemic conditions.¹⁷ I will argue that this type of case generates a problematic result for the proponents of CFR (henceforth, Revenge).

The view that CP and DP are both perfect realizers of Needs is a view that is certainly available in conceptual space. It is also a *likely* scenario given the assumptions behind CFR. Here are at least two reasons for why this is so. First, CFR takes moral kinds to be *functional* in nature. And it is plausible that functional properties, *in general*, are multiply realizable.¹⁸ To illustrate, consider a non-moral functional kind like “vision”. Roughly-speaking, vision is realized in terrestrial mammals like macaques, cats, and mice by having, among other things, the same type of brain region (visual area V1). Still, experimental work in neuroimaging has demonstrated that there are important variations across species concerning the neural realizers of these psychological properties (see Aizawa & Gillett 2009: sec. 1).¹⁹ So, even under the various constraints that apply to brain activities, it is empirically plausible that psychological properties can be multiply realized.²⁰

Secondly, the view that CP and DP are perfect realizers of Needs is supported by the set-up of MTE-cases under CFR. To be clear, for CFR to take off the ground, all that is required is that CP and DP are realizers of Needs (either perfect or imperfect). However, for CFR to be a *dialectically powerful* strategy, it must be the case that CP and DP both fulfil the Needs-role fairly well. But if CP and DP, at the very least, implement the Needs-role *well*, then the proponent of CFR cannot just stipulate that one of the two realizer-types is perfect whereas the other is not. In other words, it is not clear in what sense the scenario where CP is a perfect realizer of Needs whereas DP is not, is more plausible than the scenario that *both* CP and DP

¹⁶ Disagreement of this sort is routinely found in the history of the sciences. For discussion in the context of a causal theory of reference see Psillos (1999: ch. 12).

¹⁷ For this reason, I disagree with Väyrynen’s claim (2018: 5) that a scenario involving divergence alongside referential stability involves an epistemically unclear situation.

¹⁸ The relevant literature is vast. For both conceptual and empirical arguments for the multiple realizability of a variety of non-moral functional kinds see Aizawa & Gillett (2009; see also 2011). For recent doubts on the motivation of multiple realization via appeal to the neuroplasticity of the brain see Maimon & Hemmo (2022) (for discussion see Bickle 2020).

¹⁹ It should also be stressed that the relevant realizers here are *perfect*: there is no question about whether visual activity in macaques is ‘better’ implemented compared to visual activity in cats.

²⁰ The same point can be made more forcefully by appealing to more quixotic perfect realizers like ectoplasmic states, etc. (I return to the question of what constraints apply to multiple realization in section 5).

are perfect realizers of Needs.²¹ The proponent of CFR should tell us about the relevant constraints that are at play and whether those constraints are strong enough to single out CP or DP as the perfect realizer of Needs (more on this in section 5). Under these considerations those who adopt CFR should take the possibility of Revenge seriously.

4. Clarifying Revenge

Before trying to spell out why Revenge is a threat to the proponent of CFR, I will highlight what kind of problem Revenge is *not*. This is worth specifying given that Revenge may be conflated with superficially similar problems. First, Revenge is not an instance of the so-called Qua problem (cf. Rubin 2008: 325; McPherson 2013: 132-3) (for the Qua problem see Devitt & Sterelny [1999]). The Qua problem is a problem traditionally proposed against causal theories of reference. Roughly, the idea is this: If reference is fixed purely in terms of causation, then it is unclear how agents can accurately pinpoint to what they are referring to. The reason for this is that the causal relation is more coarse-grained than our referential intentions. For example, if I successfully refer to a cat, this cannot be purely in virtue of my causal connection to that cat. The reason for this is that there is an equally eligible referent in the vicinity, call it cat* (a cat with some particles missing from its tail), that the relevant causal relation cannot properly differentiate.

What the two challenges have in common is that they both concern a case of referential divergence. To see the difference, however, note that they can come apart. Even if there is no more than one perfect realizer of Needs, it could still be the case that a purely causal connection will not suffice for successful reference. In such a case, however, the two communities are still regulated by the same entity. It is just that the referential relation improperly latches onto whatever is intended by the two communities. Analogously, a solution to the Qua problem does little to tackle Revenge. Even if we somehow find a way to successfully pinpoint the relevant referent, it might still be the case that there is a different perfect realizer of Needs that could also play the same role.²²

Neither is Revenge a problem concerning the level of grain of the functional profile of moral *terms*. H&T argue that those who take moral terms to be individuated functionally are faced with a dilemma (1996; 2000; see also Rubin 2008: sec. 7.3.). According to the first horn, the characterization of the relevant functional role is too *weak*. For example, it could be that Earthlings and T-Earthlings use moral terms simply

²¹ This conclusion also gets additional support from the observation that there are first-order moral theories which, on their own, are plausible (recall rule-consequentialism and Rossian deontology from the previous section).

²² I find Boyd's solution to the Qua problem very convincing (2012: 53). Still, I will be focusing on Revenge as, in contrast to the Qua problem, it is a problem tailored to CFR and CM, not just the causal theory of reference in general. As Copp (1995: 235-6) also notes, if the Qua problem is a real threat, then it applies to a variety of cases (not just the moral case).

“to evaluate persons, institutions, and so forth” (H&T 2013: 2-3).²³ In turn, this means that there will be *too many* referents in a way that results to what H&T call *standard relativism*: Earthlings and T-Earthlings will refer to different entities and thus there will not be any genuine moral disagreement. To see this, note that if the relevant functional role is characterized thinly (as described above) then there is no substantive sense in which the two communities are regulated by the *same* referent.²⁴

According to the second horn, however, the characterization of the relevant functional role is *too strong*. For example, it could be argued that moral terms (used by Earthlings) are functionally individuated according to some consequentialist first-order theory. But then we get what H&T call *chauvinistic conceptual relativism*: It would be chauvinistic to say that only Earthlings refer to moral goodness given that terms used by T-Earthlings could be individuated in terms of some (say) deontological first-order theory. Again, the problematic result is that Earthlings and T-Earthlings do not refer to the same entity and thus cannot be in genuine moral disagreement.

Revenge is neither of these two challenges. To see this, notice that CFR can be used *as a response* to both horns of H&T’s dilemma. Consider the first horn. According to CFR, it is simply not true that the functional role of moral terms should be weakly individuated. As noted, there are many plausible candidates for a substantive characterization of such a role. In this paper I focus on Needs but, as noted, there are other accounts that could also fit the bill (e.g. societal stability as per Copp [2007]). Note, also, that individuating the function of moral terms according to Needs does not commit the proponent of CFR to adopt the second horn of H&T’s dilemma. Countenancing Needs does not, by itself, entail anything substantive about first-order moral theory. It is true that Boyd adopts a consequentialist reading of Needs (i.e. what he calls homeostatic consequentialism [1988: sec. 4.3.]) but this is purely for illustration purposes. For all we know, Needs is compatible with a deontological theory according to which adhering to some deontic constraints satisfies some appropriately defined cluster of human needs (see Boyd 1988: 238, 340).

So, CFR can avoid both horns of H&T’s dilemma: Earthlings and T-Earthlings are in genuine disagreement since their respective terms are regulated by the same functional role (say, the satisfaction of Needs) but not in a way where a particular first-order moral theory is singled out (thus avoiding chauvinistic conceptual relativism). But Revenge remains, nonetheless. It could still be the case that Needs have more

²³ In other work, H&T (1996: 27-8, sec. 7) cash out this “weak” characterization of moral terms by appealing to a folk, *common-sense*, theory of moral discourse. However, this move is potentially dialectically weak given that there are sophisticated theories of folk moral functionalism which, arguably, *would not* be weak in H&T’s sense (cf. Jackson 1998: ch. 6).

²⁴ Rather, commonality of reference is secured in a trivial way. To compare, consider how according to a simplistic causal theory of reference the terms ‘phlogiston’ and ‘oxygen’ have the same referent if the only constraint at play is our referential intention to refer to the phenomenon underlying the uses of those terms (cf. Psillos 1999: ch. 12). Similarly, Kraemer’s (1991: 469) suggestion that the functional role could be characterized in terms of what is *minimally necessary* in order to define morality, also runs into the same problematic result.

than one perfect realizer.²⁵ In the next section I argue that this is a serious problem for the proponent of CFR.

5. The threat of Revenge

To take stock, it is worth stating Revenge once more:

- (1) CFR entails that CP and DP are both realizers of a higher-order, functional, property like Needs.
- (2) It is plausible that CP and DP (and perhaps additional property-types) are *perfect* realizers of Needs (this is what I have called Revenge).
- (3) So, it is plausible that CFR entails Revenge.

In this section, I will argue that Revenge is a threat worth taking seriously. Specifically, I will argue that the proponent of CFR can tackle Revenge in two ways: either by biting the bullet and directly countenancing Revenge, or by attacking (2) (i.e. the claim that there is more than one perfect realizer-type of Needs).

But the acceptance of either horn constitutes a serious cost. Specifically, the following dilemma is generated: either CFR entails a form of value pluralism that is both controversial and contrary to the spirit of moral realism, or CFR avoids such value pluralism but only by making substantive metaphysical/empirical claims. I will consider each horn in turn.

5.1. Biting the bullet

According to (2), it is plausible that CP and DP are both perfect realizers of Needs. At first glance, this might not seem like a problematic result. If CP and DP are both perfect realizers of Needs, then the proponent of CFR would have to countenance a moderately pluralistic version of moral realism (i.e. one according to which goodness is, fundamentally, brought about by either CP or DP) (see Wong 2006). But, perhaps, the moral realist can tolerate this kind of pluralism.

In response, it should be noted that even moderate pluralistic versions of moral realism have been challenged in the literature (e.g. Li 2019). More importantly, however, it is not clear that Revenge results in a *moderately* pluralistic version of moral realism. Note that nothing in CFR prevents Needs from having *more* than two perfect realizers. For all we know, the number of perfect realizers of Needs could be abundant.

²⁵ Note that I have framed Revenge in terms of the first-order divergence and second-order, functional, convergence between the two communities, rather than in terms of their putative disagreement. To see why, consider a scenario where an Earthling and a T-Earthling, after realizing that Revenge is the case, decide to update their beliefs in a way that reflects that Needs is realized by the disjunction of CP and DP (to compare: if Water is a functional type, then presumably Earthling and T-Earthlings could take Water to refer to the disjunction between H₂O and XYZ). Even if we assume that their first-order disagreement is resolved at this point, as I will argue in the next section, the threat of Revenge remains unchanged (see section 5). [Acknowledgments].

To be clear, moral realism is *compatible* with there being many perfect realizers of goodness insofar as these realizers are broadly stance-independent entities. So, where does the problem lie?

First, value pluralism (moderate or not) is philosophically parochial and controversial (see Heathwood 2015). Revenge seems to imply that there are many radically different first-order moral theories which are all true at the same time. But, ideally, we would not want for CFR to entail anything substantive about first-order moral theory. And, more importantly, if CFR *were* to involve substantive first-order commitments, we would not want for these commitments to be controversial. This last point is even more pressing if the resulting kind of moral pluralism is not moderate. Recall (from section 3) that it is *prima facie* plausible that there are *many* perfect realizers of Needs.

To be clear, I do not wish to suggest that such first-order commitments can be used as a *decisive* argument against CFR. After all, it could be that a kind of moral realism that adopts CFR could have other (non-moral) virtues which would outweigh the controversial nature of these first-order commitments. Rather, I am simply noting that, at any case, these commitments count *against* CFR. This is compatible with the philosophical constraints that moral realists typically apply to their own theory. For example, realist moral naturalists wish to accommodate moral practice and moral phenomenology (e.g. Railton 1986: 171; Jackson 1998: ch. 5) even if, as mentioned, there are non-moral considerations that pull from the opposite direction.

Secondly, the resulting picture is suspiciously close to what typical relativist accounts take moral reality to be. Roughly, the moral relativist countenances the possibility of distinct communities with different first-order moral views, all being true relative to each community's set of pure moral principles. The kind of moral pluralism I am considering is compatible with this picture, *minus* the relativization component.²⁶ In essence, Revenge is compatible with the possibility of such communities all being correct in their moral judgements since every community has presumably *tracked* a perfect realizer of Needs. But

²⁶ This sort of relativism should be distinguished from the kind of “innocent” relativism that naturalist moral realists typically accept. It is true that, for example, Boyd characterizes the relevant notion of needs in terms of various “nonmoral factual” factors (e.g. physical and psychological characteristics of human beings) (Boyd 1988: 347) and Copp relativizes societal stability in terms of the circumstances of each society (cf. Copp 2007: 244, note 25). But that's not the kind of view that's typically labeled as “moral relativism” (in this sense, “relationalism” should be distinguished from “relativism”). As it will become apparent given my upcoming remarks, what is at stake here concerns the number of perfect realizers.

this means that the resulting kind of moral realism is extensionally equivalent to moral relativism.²⁷ This is surely not the result that most moral realists are looking for.²⁸

5.2. Restricting multiple realization

There is a way to resist the claim that Revenge involves a commitment to value pluralism. After all, it is unlikely that multiple realization occurs unrestrictedly (or, to echo Aizawa & Gillett [2009], multiple realization is not ‘ghostly’). This is in virtue of what Ned Block calls the *Disney Principle* (1997: 120-4). According to this principle, multiple realization always operates under some nomic constraints. There are conceivable realizers of S which, nonetheless, cannot *actually* realize S. At least some of its putative realizers will be excluded in virtue of the laws that operate in the relevant world. Minds cannot be realized by collections of blue cheese since, given the nature of blue cheese and the laws that govern dairy-like substances, the relevant function is impossible to implement. Perhaps the same principle can be appealed to in the case of Needs. In such a scenario, DP might be conceivable as a realizer of Needs but impossible to instantiate. Or, perhaps, instantiations of DP are nomologically possible but given the nomic constraints that apply to Needs *themselves*, DP are not among its genuine realizers.²⁹

I do not think that one can respond to Revenge simply by appealing to the Disney Principle. After all, the Disney Principle merely states that there *are* constraints at play, while staying neutral concerning the nature of those constraints. For all we know, the constraints governing the multiple realizability of moral facts are extremely weak. To illustrate, note that even moral antirealists appeal to minimal constraints such

²⁷ To be more precise, in such a case moral realism would be extensionally equivalent but *hyperintensionally* distinct from moral relativism: the *grounds* of moral goodness under moral realism are different than its grounds under moral relativism. In the former case, moral goodness is grounded by stance-independent natural facts, whereas in the latter case it is grounded by the contingent values and commitments of various societies. So, even though both moral realism and moral relativism accept the same moral truths (in this scenario), they still countenance different grounds for these truths. [Acknowledgments].

²⁸ The pluralist account that I have sketched is similar to what Harman (1978) calls *meta-ethical relativism*. Roughly, this version of moral relativism allows for conflicting judgments about a particular case to be all correct. In fact, I want to suggest that by appealing to something like Needs it could be said that meta-ethical relativism is *strengthened*. Initially, the claim that two conflicting moral judgments can be both correct, seems to entail some kind of inconsistency (Harman 1978: sec. 7). But there is no inconsistency if we take the relevant regulators of these judgments to be distinct perfect realizers of (say) Needs. In this sense, the two communities are both correct in their judgments in the sense that they have tracked two distinct ways of realizing the *same* higher-order function.

²⁹ The Disney Principle is independently plausible. Specifically, it is a direct consequence of the view that special science laws supervene upon the more fundamental laws of nature. Now note that typical proponents of CM propose the realization-fact ‘CP functionally realize Needs’ as a special science law (e.g. Sturgeon 1985). In this sense, it is expected that the multiple realization thesis of the form ‘{CP or DP or ...} functionally realize Needs’ is subject to nomic constraints from more fundamental laws. A clear example is the following: a putative perfect realizer of Needs that violates the law of the conservation of energy is nomologically impossible.

as coherence (e.g. Harman 2014: 6; Street 2006: sec. 10; Capps et al. 2007). In this sense, even under such accounts it will not be the case that *just about anything* could be a genuine perfect realizer of moral goodness.³⁰

Of course, the relevant constraints might be strict *enough* in a way that the resulting picture would vindicate typical first-order assumptions about moral theory. For example, Copp (1995) argues that goodness depends on the rational requirements of a given society. This significantly narrows down the relevant set of options. Given that *every* society has a common set of fundamental needs, many types of moral codes will be excluded (Wong [2006] also appeals to human nature to make a similar point).³¹

Alternatively, one could claim that there are constraints from the domain of first-order moral theory. Plausibly, not just ‘anything’ can be a moral theory. The history of ethics illustrates that ethical theories usually appeal to one of the following types of phenomena. For example: the mental states of an idealized agent (as per the Ideal Observer theory), the mental states of an ideal/divine being (as per Divine Command theory), the consequences of human actions (as per Consequentialism), certain deontic constraints (as per Deontology), the actual or hypothetical contracts between agents (as per Contractualism), etc. So, perhaps, the above is enough to narrow down the set of perfect realizers of Needs.³²

I have two responses. First, similarly to the Disney Principle, I am not optimistic that these constraints are strong enough. To see this, consider that each of these constraints can generate many *types* of moral theories (e.g. there are many types of consequentialist theories, such as rule-utilitarianism, or Boyd’s homeostatic consequentialism). So, even if the list above is exhaustive (which is doubtful), it is still the case that the history of ethics generates a multitude of prima facie plausible ethical theories (thus partially supporting Revenge). Secondly, if one were to narrow down the set of perfect realizers by claiming that a given moral theory is correct (say, Kantian Deontology), then that would constitute a controversial first-order moral assumption.

³⁰ One might be tempted to apply a *naturalness* constraint in order to single out a particular perfect realizer. Presumably, one of the putative realizers of Needs will be more *natural* (in the Lewis [1983] sense) than its competitors. But in whichever way one cashes out the relevant notion of naturalness, the perfect realizers at play will score the same degree of naturalness. On the classic interpretation, a property is natural when it is defined in terms of other natural properties and its definition is relatively short. But this interpretation is notoriously problematic when it is applied to phenomena that are quite distant from the fundamental level (cf. Dunaway & McPherson 2016: sec. 3.4.). On a context-relative account of naturalness, on the other hand, the relevant perfect realizers all turn out to be *equally* natural (cf. Taylor 2016). This is because the context at hand would surely concern the implementation of the appropriate function. But it is stipulated that perfect realizers fulfil that role in the fullest sense. For a similar problem involving ties in the context of naturalness see Mokriski (2020: 22).

³¹ It is worth emphasizing that the severity of Revenge concerns the possibility of multiple *perfect* realizers of (say) Needs. Not just any multiple realization thesis will motivate Revenge. This is particularly important since some multiple realization theses are trivial and dialectically unimportant (e.g. the realization of a mental state by two similar tokens of the same neural type) (cf. Polger & Shapiro 2016).

³² [Acknowledgments].

So, it seems that whether the constraints that apply to Needs are strict enough is ultimately an empirical issue or an issue of first-order moral theorizing. But this doesn't make Revenge any less pressing as a problem. In fact, noting that the proponent of CFR needs to appeal to substantive empirical/metaphysical or normative constraints to narrow down the relevant set of perfect realizers (in a way that delivers a desirable form of moral realism) is *precisely* why Revenge is a problem that should be taken very seriously. The reason for this is that whether there *are* indeed such constraints is a highly contested issue.

In saying this I do not wish to downplay the various proposals that moral realists have made in the literature to that effect. Copp (as mentioned) appeals to rational requirements (and also to the notion of a *global* society [2007: 24-5]), Wong (2006) appeals to human nature, Railton construes moral goodness in terms of what is good for an *ideal* agent (Railton 1986), and Jackson (1999) is optimistic that a *mature* folk moral theory will plausibly indicate the existence of these constraints. But, and here is the crucial point, *none* of the above proposals is obviously true. Adopting any of them amounts to a substantive theoretical commitment.

After all, there are well-known positions in the literature that go against these views. It is a datum that there is *prima facie* irresolvable moral disagreement. And the fact that there are many perfect realizers of Needs would provide a clear-cut explanation for the apparent persistence of such disagreement. The typical moral realist response is to try to explain away the phenomenon by using some kind of reconstruction of the relevant cases. For example, the fact that two communities are in moral disagreement about (say) wealth distribution could be fundamentally explained by the fact that each community is governed by different contingent societal circumstances (e.g. one community occupies a position of scarcity) (see Brink 1984: sec. 3). But, again, note that these strategies will not work in the case of Revenge since, *ex hypothesi*, the relevant realizers are *perfect*.

The above illustrates that CFR cannot, *on its own*, fully discharge Revenge and hence solve the Moral Twin Earth problem. As noted, the proponent of CFR faces a dilemma at this point. On the one hand, CFR, has controversial implications about first-order moral theory. As mentioned, it entails a kind of (at best moderate) value pluralism that is incompatible with how moral realists typically understand their view. Some might be ready to bite this bullet (given other purported non-moral theoretical gains), but I take it that this is not something that everyone is ready to accept. On the other hand, it could be that CFR is coupled with various metaphysical/empirical assumptions in the hopes of narrowing down the set of possible perfect realizers.³³ As noted, these assumptions will be very substantive and philosophically parochial. More importantly, however, it is not at all obvious that these assumptions are going to end up

³³ The most extreme case of this strategy is the complete rejection of multiple realization as a general phenomenon (for recent attacks see Bickle 2020: secs. 2, 3.2.).

being *true*. In this sense, any way you cut it, the proponent of CFR needs to address Revenge and its implications.³⁴

6. Invitation for future work

I have argued that Revenge is a problem that warrants serious consideration. Still, I have also argued that we will not know what Revenge *really* entails until we know more about the nature of Needs (or the functional role of one's choice). To echo Boyd, the question of identifying the relevant set of important human needs and providing fine-grained information about their nature "is a potentially difficult and complex empirical question" (1988: 329).³⁵ It is striking, however, that there is very little substantive literature on these questions.³⁶ For this reason, it is very important for the proponent of CFR to not simply appeal to some abstract notion of Needs but also provide a *theory* of Needs.

Acknowledgments

[Removed]

References

- Adams, Robert Merrihew (1999). *Finite and Infinite Goods: A Framework for Ethics*. Oxford University Press.
- Aizawa, Kenneth & Gillett, Carl (2009). The (Multiple) Realization of Psychological and Other Properties in the Sciences. *Mind & Language*, 24(2): 181–208.
- Aizawa, Kenneth & Gillett, Carl (2011). "The Autonomy of Psychology in the Age of Neuroscience", in *Causality in the Sciences*, Phyllis McKay Illari, Federica Russo, and Jon Williamson (eds.), Oxford: Oxford University Press, 202–223.
- Baysan, Umut (2015). Realization Relations in Metaphysics. *Minds and Machines* (3):1-14.
- Bickle, John (2020). "Multiple Realizability" in *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2020/entries/multiple-realizability/>>.
- Block, Ned (1997). Anti-Reductionism Slaps Back. *Noûs* 31 (s11):107-132.

³⁴ I make no claims about whether the two horns of this dilemma are equally threatening. For my purposes, it suffices to say that they are both worthy of attention.

³⁵ To compare, consider the initial MTE problem where the mere conceptual possibility of a twin Earth is *prima facie* problematic.

³⁶ Although, see Brock (2019). At any rate, the current discussion on the notion of needs mainly focuses on their normative import, rather than their metaphysics. One notable exception here is arguably Copp (1995: ch. 10).

- Boyd, Richard (1988). How to be a Moral Realist. In G. Sayre-McCord (ed.), *Essays on Moral Realism*. Cornell University Press. pp. 181-228.
- Boyd, Richard (2003a). Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, part I. *Philosophy and Phenomenological Research* 66 (3):505–553.
- Boyd, Richard (2003b). Finite beings, finite goods: The semantics, metaphysics and ethics of naturalist consequentialism, part II. *Philosophy and Phenomenological Research* 67 (1):24–47.
- Boyd, Richard (2012). What of pragmatism with the world here? In Maria Baghramian (ed.), *Reading Putnam*. Routledge. pp. 39.
- Braddon-Mitchell, David & Nola, Robert (2009). Introducing the Canberra Plan. In David Braddon-Mitchell & Robert Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. MIT Press. pp. 1--20.
- Brink, David O. (1984). Moral realism and the sceptical arguments from disagreement and queerness. *Australasian Journal of Philosophy* 62 (2):111 – 125.
- Brink, David O. (2001). Realism, Naturalism, and Moral Semantics. *Social Philosophy and Policy* 18 (2):154.
- Brink, David Owen (1989). *Moral Realism and the Foundations of Ethics*. Cambridge University Press.
- Brock, Gillian and David Miller, "Needs in Moral and Political Philosophy", The Stanford Encyclopedia of Philosophy (Summer 2019 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/sum2019/entries/needs/>>.
- Capps, David; Lynch, Michael P. & Massey, Daniel (2009). A coherent moral relativism. *Synthese* 166 (2):413 - 430.
- Chalmers, David J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- Copp, David (1995). *Morality, Normativity, and Society*. Oxford University Press.
- Copp, David (2000). Milk, honey, and the good life on moral twin earth. *Synthese* 124 (1-2):113-137.
- Copp, David (2007). *Morality in a Natural World: Selected Essays in Metaethics*. Cambridge University Press.
- Dowell, Janice L. (2016). The Metaethical Insignificance of Moral Twin Earth. *Oxford Studies in Metaethics* 11.
- Dunaway, Billy & McPherson, Tristram (2016). Reference Magnetism as a Solution to the Moral Twin Earth Problem. *Ergo: An Open Access Journal of Philosophy* 3.
- Freiman, Christopher (2014). Goodness and Moral Twin Earth. *Erkenntnis* 79 (2):445-460.
- Harman, Gilbert (1978). What is moral relativism? In A. I. Goldman & I. Kim (eds.), *Values and Morals*. Boston: D. Reidel. pp. 143--161.
- Harman, Gilbert (2015). Moral relativism is moral realism. *Philosophical Studies* 172 (4):855-863.

- Heathwood, Chris (2015). Monism and Pluralism about Value. In Iwao Hirose & Jonas Olson (eds.), *Oxford Handbook of Value Theory*. Oxford University Press. pp. 136-157.
- Horgan, Terence & Timmons, Mark (1992). Troubles on moral twin earth: Moral queerness revived. *Synthese* 92 (2):221 - 260.
- Horgan, Terence & Timmons, Mark (1996). From Moral Realism to Moral Relativism in One Easy Step. *Critica* 28 (83):3-39.
- Horgan, Terence & Timmons, Mark (2000). Copping out on moral twin earth. *Synthese* 124 (1-2):139-152.
- Horgan, Terence & Timmons, Mark (2013). Twin Earth, Moral. In International Encyclopedia of Ethics, H. LaFollette (Ed.). doi:[10.1002/9781444367072.wbiee506](https://doi.org/10.1002/9781444367072.wbiee506)
- Jackson, Frank (1998). *From Metaphysics to Ethics: A Defence of Conceptual Analysis*. Oxford University Press.
- Kraemer, Eric Russert (1991). On The Moral Twin-Earth Challenge to New-Wave Moral Realism. *Journal of Philosophical Research* 16:467-472.
- Laurence, Stephen; Margolis, Eric & Dawson, Angus (1999). Moral Realism and Twin Earth. *Facta Philosophica* 1:135-165.
- Levy, Neil (2011). Moore on Twin Earth. *Erkenntnis* 75 (1):137-146.
- Lewis, David (1983). New work for a theory of universals. *Australasian Journal of Philosophy* 61 (4):343-377.
- Li, Yong (2019). Moral Ambivalence: Relativism or Pluralism? *Acta Analytica* 34 (4):473-491.
- Maimon, Amber & Hemmo, Meir (2022). Does Neuroplasticity Support the Hypothesis of Multiple Realizability? *Philosophy of Science* 89 (1):107-127.
- McPherson, Tristram (2013). Semantic Challenges to Normative Realism. *Philosophy Compass* 8 (2):126-136.
- Merli, David (2002). Return to Moral Twin Earth. *Canadian Journal of Philosophy* 32 (2):207-240.
- Michael Devitt, & Kim Sterelny, (1999). *Language and Reality: An Introduction to the Philosophy of Language*. Wiley.
- Mokriski, David (2020). The Methodological Implications of Reference Magnetism on Moral Twin Earth. *Metaphilosophy* 51 (5):702-726.
- Plunkett, David & Sundell, Timothy (2013). Disagreement and the Semantics of Normative and Evaluative Terms. *Philosophers' Imprint* 13 (23):1-37.
- Polger, Thomas W. & Shapiro, Lawrence A. (2016). *The Multiple Realization Book*. Oxford University Press UK.
- Psillos, Stathis (1999). *Scientific Realism: How Science Tracks Truth*. Routledge.
- Putnam, Hilary (1975). The meaning of 'meaning'. *Minnesota Studies in the Philosophy of Science* 7:131-193.

- Railton, Peter (1986). Moral realism. *Philosophical Review* 95 (2):163-207.
- Railton, Peter (1993). Noncognitivism about rationality: Benefits, costs, and an alternative. *Philosophical Issues* 4:36-51.
- Rubin, Michael (2008). Sound intuitions on Moral Twin Earth. *Philosophical Studies* 139 (3):307-327.
- Shoemaker, Sydney (2007). *Physical Realization*. Oxford University Press UK.
- Sonderholm, Jorn (2013). Unreliable Intuitions: A New Reply to the Moral Twin-Earth Argument. *Theoria* 79 (1):76-88.
- Street, Sharon (2006). A Darwinian dilemma for realist theories of value. *Philosophical Studies* 127 (1):109-166.
- Sturgeon, Nicholas L. (1985). Moral Explanations. In James Rachels (ed.), *Ethical Theory 1: The Question of Objectivity*. Oxford University Press.
- Taylor, Elanor (2016). Naturalness in Context. *Inquiry: An Interdisciplinary Journal of Philosophy* 59 (4):1-24.
- Väyrynen, Pekka (2018). A Simple Escape from Moral Twin Earth. *Thought: A Journal of Philosophy* 7 (2):109-118.
- Wisdom, Jeffrey (2021). A dose of reality for moral twin earth. *Philosophical Psychology* 34 (6):784-804.
- Wong, David B. (2006). *Natural Moralities: A Defense of Pluralistic Relativism*. Oxford University Press.