

# MODEL THEORY, HUME'S DICTUM, AND THE PRIORITY OF ETHICAL THEORY

JACK WOODS

*University of Leeds*

BARRY MAGUIRE

*University of North Carolina, Chapel Hill*

**Abstract:** *It is regrettably common for theorists to attempt to characterize the Humean dictum that one can't get an 'ought' from an 'is' just in broadly logical terms. We here address an important new class of such approaches which appeal to model-theoretic machinery. Our complaint about these recent attempts is that they interfere with substantive debates about the nature of the ethical. This problem, developed in detail for Daniel Singer's and Gillian Russell and Greg Restall's accounts of Hume's dictum, is of a general type arising for the use of model-theoretic structures in cashing out substantive philosophical claims: the question of whether an abstract model-theoretic structure successfully interprets something often involves taking a stand on non-trivial issues surrounding the thing. In the particular case of Hume's dictum, given reasonable conceptual or metaphysical claims about the ethical, Singer's and Russell and Restall's accounts treat obviously ethical claims as descriptive and vice versa. Consequently, their model-theoretic characterizations of Hume's dictum are not metaethically neutral. This encourages skepticism about whether model-theoretic machinery suffices to provide an illuminating distinction between the ethical and the descriptive.*

## 1. Introduction: Hume's Dictum

... as this ought, or ought not, expresses some new relation or affirmation, 'tis necessary that it shou'd be observ'd and *explain'd*; and at the same time that *a reason should be given*, for what seems altogether inconceivable, how this new relation can be a deduction from others, which are entirely different from it ... [I] am persuaded, that a small attention [to this point] wou'd subvert all the vulgar systems of morality ... (Hume, *Treatise on Human Nature*, 3.1.1)

---

**Contact:** Jack Woods <j.e.woods@gmail.com>, Barry Maguire <barrymaguire@gmail.com>

*Hume's dictum*, the idea that that we cannot get an ethical claim from a descriptive one, has exercised much influence over theorists in ethics and meta-ethics.<sup>1</sup> Somewhat unfortunately, the simplest logical characterization of Hume's dictum—no descriptive proposition entails an ethical proposition—is demonstrably false. A.N. Prior (1960) showed this with several simple counterexamples. Let  $N$  be an ethical sentence, say, that it is impermissible to starve the Irish, and let  $M$  be a descriptive claim, say, that most people cannot digest grass. Presumably,  $\neg M$  is also descriptive. If  $N \vee M$  is descriptive, then the descriptive  $\neg M$  together with  $N \vee M$  entails the ethical  $N$ . If  $N \vee M$  is ethical, then it is entailed by the descriptive  $M$ . Assuming  $N \vee M$  is either ethical or descriptive, we have a decisive argument that the descriptive can entail the ethical.<sup>2</sup> These entailments are impeccable; in the ordinary sense of entailment, we really do have a case of the ethical being entailed by the descriptive.

Prior's counterexamples laid down a gauntlet for theorists drawn to a version of Hume's dictum. A generation of theorists attempted to characterize the intuitive thesis with increasingly sophisticated logical versions of Hume's dictum.<sup>3</sup> One class of attempts treats the is-ought gap as an instance of a more general property of logical conservativeness—anytime we have a Prior-type entailment like the above, the conclusion is *ethically irrelevant* in a particular sense they define (Pigden 1989; Schurz 1997). We largely set these attempts aside.<sup>4</sup>

Our focus here is on an even more recent class of attempts; these interpret the is-ought gap in terms of formal properties of semantical frameworks. Developed by Daniel Singer (2015) and jointly by Gillian Russell and Greg Restall (2010), these accounts start with a model-theoretic characterization of what it is to be an ethical sentence (roughly in terms of some form of invariance under change in the ethical component of the semantics). They then go on to prove that, so characterized, no ethical claims are entailed by fully descriptive premises. These approaches aim to both distinguish descriptive and ethical sentences and prove that descriptive sentences do not entail ethical sentences, by appeal to just model-theoretic machinery.

Our complaint about these recent attempts is that they interfere with substantive debates about the nature of the ethical. This problem, here developed in de-

---

1. We will use 'ethical' and 'descriptive' instead of 'normative' and 'non-normative' or 'ethical' and 'non-ethical' in characterizing Hume's dictum so as to avoid worries about the broad nature of normativity, permit that some claims might be both descriptive and ethical, and connect with most literature on Hume's dictum. Clearly our point is more general, though we leave it to the reader to make the relevant adjustments to our arguments.

2. A trivial modification demonstrates that the ethical can also entail the descriptive.

3. The *locus classicus* for these discussions is Pigden (2010a), especially the contributions from Pigden and Schurz. For some further critique, see Karmo (1988), Brown (2014), and Maguire (2015).

4. For critical discussion, see Maguire (2015).

tail for Singer's and Russell and Restall's account of Hume's dictum, is of a general type arising for the use of model-theoretic structures in cashing out substantive philosophical claims: the question of whether an abstract model-theoretic structure successfully interprets something often involves taking a stand on non-trivial issues surrounding that thing. In this particular case, the problem is that in the presence of reasonable conceptual or metaphysical claims about the ethical, the taxonomies Singer's and Russell and Restall's accounts are built on treats obviously ethical claims as descriptive and conversely. These incorrectly taxonomized claims are not bizarre speculations involving ethical concepts; they include such banalities as the claim that the permissible is possible and that pain is somewhat disvaluable. This taxonomical problem is present even when the putative conceptual truths, such as those just listed, cannot plausibly be seen as interfering with the intuitive taxonomical distinction between the ethical and the descriptive.

This issue creates further problems with their characterization of Hume's dictum. As we will shortly show, views like John Searle's (1964) which deliberately violate Hume's dictum by analytically linking descriptive facts with ethical facts are treated as views which accept the is-ought gap. This since views like Singer's define the ethical and the descriptive in such a way that the intuitively descriptive premises of Searle's is-ought inference are ethical. This strains belief.

At a minimum, someone can accept Singers's and Russell and Restall's accounts of Hume's dictum only if they deny there are conceptual or metaphysically necessary truths about the ethical of the sort we will shortly consider. But we view these putative conceptual truths as claims which should be arbitrated by ethical, metaethical, and linguistic considerations. In particular, how we should represent notions like 'ought,' 'wrong,' or 'permitted' formally, including which conceptual truths we hard-wire into the semantics, is itself a matter for substantive theorizing by ethicists, metaethicists, and linguists.

Since these characterizations of Hume's dictum are only plausible when they reject at the outset many philosophically plausible metaethical positions, they are problematically partisan.<sup>5</sup> They are also taxonomically implausible since the difference between the ethical and the descriptive should not depend on, for instance, whether it is a conceptual truth that all permissible things are possible.<sup>6</sup> Finally, they are incapable of representing a range of metaethical views like

---

5. Here, as in our title, we echo an earlier point of Geoff Sayre-McCord's: "Deontic logic cannot hover above the maelstrom of moral argumentation. As a result, if we demand neutrality we must abandon deontic logic. Conversely, if we embrace deontic logic, we must allow our "logic" substantive (and controversial) theorems" (1986: 179). Compare Schurz (2010a: 211).

6. In contrast, the aforementioned approaches of Pigden and Schurz take no particular stand on the taxonomical division between ethical and descriptive claims. They show, roughly, that certain 'ought'-containing sentences are not entailed by collections of premises which do not, themselves, contain 'ought' claims. Because they take no particular stand on the

Searle's which intuitively and intentionally reject the is-ought gap. In short, it's far from clear that the distinction between the ethical and the descriptive admits of an illuminating model-theoretic characterization and it's clear that the model-theoretic characterizations given by Singer and Russell and Restall are not metaethically neutral. We now turn to developing these criticisms in detail, starting with Singer (2015), then addressing Russell and Restall (2010).

## 2. Singer's Account

### 2.1. *Singer's Approach: NORM INVARIANCE and WORLD-NORM GAP*

Singer's discussion (2015) starts with an intuitive shortcoming with Prior's counterexamples: there is no way to argue from the premises of a Prior example to the conclusion without already being a position to assert the conclusion. Intuitively, this means that Prior's arguments are lousy ways of acquiring information about the ethical even though they're valid. Singer points out that any time we can reason from intuitively descriptive premises to an intuitively ethical conclusion, this is due to the fact that the possibilities consistent with the premises already validate the conclusion. In other words, we only get from the conclusions of these arguments ethical content that was already part of our starting point. He then argues that if the premises of an argument are insensitive to ethical facts (in his terminology, if they are *norm invariant*), then the conjunction of the premises and the conclusion is also insensitive to the ethical facts. This means that we acquire no new ethical information by so arguing.

Singer's point here is absolutely apt—you can see why the Prior arguments don't show that we can really obtain an 'ought'-fact from an 'is'-fact by noticing that anytime we can reason from the latter to the former, we are already in a position to accept the former. This follows from the fact that deductive argumentation is non-ampliative; we never gain new information, in the sense of narrowing the set of (ethical and factual) possibilities consistent with our beliefs, by drawing deductively valid conclusions. Singer's reply to Prior is thus that if

---

taxonomical division between ethical and descriptive claims, their results hold for broad classes of expressions, including obviously ethical ones like 'ought', but also including expressions having little to nothing to do with ethicality, such as epistemic operators. Likewise, their results do not rule out the existence of conceptual truths linking the descriptive with the ethical, though the extent of certain gap theses, like Schurz's, depends on whether or not there are such truths. Given this, the results of Pigden and Schurz do not interfere with substantive debates about the nature of the ethical. Pigden and Schurz's views also fail to fully capture the extent of Hume's dictum, at least as currently developed, but that's life. We have no substantive complaints about their results; we find them interesting, though they are largely not particular to the ethical. We think there is important work to be done developing Pigden and Schurz's approaches: in particular, in exploring exactly which "substantive" conclusions can be entailed by which.

you start with a set of premises that don't make an ethical difference, then you will never shake loose a consequence that makes an ethical difference.

To spell out his key notion of norm invariance, Singer appeals to the semantical framework Allan Gibbard uses to develop his expressivist view of ethical discourse. This semantics replaces the simple notion of a possible world with pairs consisting of worlds and sets of norms. The latter are meant to govern what we are permitted, obliged, and forbidden from doing. We can write these pairs  $\langle w, n \rangle$ . We proceed as usual for descriptive terms, treating propositions as sets of world-norm pairs, and logical operations as set-theoretic operations (intersection, union, restricted complementation) on these. Note that it isn't clear, from Singer's discussion, whether he takes the field of world-norm pairs to model *conceptual* or *metaphysical* possibility; our point holds on both construals, though the examples differ slightly. We will flag this up when it is relevant.

The ethical component of a world-norm pair figures essentially in the semantics for expressions like 'ought', 'wrong', etc. in the obvious ways. Truth at a world-norm pair, written  $\langle w, n \rangle \models \varphi$ , is defined with reference both to the worldly component and the ethical component of a world-norm pair since a sentence might contain both ethical and descriptive materials. We can intuitively think of  $\langle w, n \rangle \models \varphi$  as saying, when  $\varphi$  is an ethical claim like 'we ought not murder', something like 'the norms in  $n$  forbid murdering.' This corresponds model-theoretically, on the Gibbardian approach, to the claim that  $\langle w, n \rangle$  is in the proposition (i.e., set of world-norm pairs)  $\llbracket \varphi \rrbracket$ . When  $\varphi$  is a descriptive claim like 'grass is green', we can think of  $\langle w, n \rangle \models \varphi$  as saying 'grass is green in  $w$ ' or, again in Gibbard's account, as  $\langle w, n \rangle$  is in  $\llbracket \varphi \rrbracket$ .

We define the entailment of a sentence  $\varphi$  from a set of sentences  $\Gamma$ , written  $\Gamma \models \varphi$ , in the obvious way:  $\llbracket \Gamma \rrbracket$  is a subset of  $\llbracket \varphi \rrbracket$  for  $\llbracket \Gamma \rrbracket$  the intersection of the propositions denoted by the sentences in  $\Gamma$  if there are any, the empty set otherwise.<sup>7</sup> Given this set-up, we can define norm invariance easily:

NORM INVARIANCE:  $\varphi$  is norm invariant just in case  $\forall w [\exists n \langle w, n \rangle \models \varphi \Rightarrow \forall n \langle w, n \rangle \models \varphi]$ <sup>8</sup>

This says that a sentence is norm invariant just in case, given a world, it is satisfiable at a pair of that world with the choice of some norm only if it is satisfiable by the pair of that world with the choice of any norm. This captures at least one sense in which the ethical dictates of world-norm pairs are irrelevant to whether or not a descriptive proposition holds at a world-norm pair: if it can be true, then any old norms would do.

7. We let  $\models$ , functioning variously as entailment or the world-verification relation, be disambiguated by context.

8. Alternatively,  $\forall w [\exists n \langle w, n \rangle \in \llbracket \varphi \rrbracket \Rightarrow \forall n \langle w, n \rangle \in \llbracket \varphi \rrbracket]$ . That is, given a world  $w$ , if  $\llbracket \varphi \rrbracket$  contains a pair  $\langle w, n \rangle$  for a set of norms  $n$ , then, for every set of norms  $k$ , it contains  $\langle w, k \rangle$ .

How does Singer use this machinery to characterize Hume’s dictum? Singer’s thought is that if the truth of a claim at  $\langle w, n \rangle$  depends essentially on  $n$ , it must be ethical. We test for essential dependence by seeing whether varying  $n$  varies whether the claim holds. This plausibly captures a sense in which the ethical is barricaded against, or insulated from, the descriptive. Note that the intuitive sketch just given makes no claims about the particular nature of the ethical; the ethical, whatever it might be, is supposed to be by nature insulated from the descriptive in this way. Even if we disagree about the nature and meaning of the ethical, any content which is recognizable as ethical should obey Hume’s dictum—especially for those pictures which aim to capture something like the intuitive sense in which the ethical is insulated from the descriptive.

With the notion of norm invariance defined, Singer argues that a particular principle, *WORLD-NORM GAP*, captures the essence of Hume’s dictum. *WORLD-NORM GAP* is a claim about the persistence of norm invariance across logical entailment:

*WORLD-NORM GAP*: If  $\Delta \models \varphi$  and all of  $\Delta$  are norm invariant and satisfiable, then  $\Delta \ \& \ \varphi$  (the conjunction of all the sentences in  $\Delta$  and  $\varphi$ ) is norm invariant.<sup>9</sup>

## 2.2. A General Observation about Model-Theoretic Strategies

We start by pointing out a potential limitation of this strategy as such. Separating out “worlds” from “norms” guarantees one kind of “formal autonomy”, but says nothing distinctive yet about the ethical. There is nothing special about norms in *WORLD-NORM GAP*.<sup>10</sup> Let a *world-octopus* pair be a pair of a possible world and an octopus, letting the octopus slot interact with some select expressions, like ‘Octo’—an expression dedicated to referring to the favoured octopus—in the obvious way. This allows us to say things like ‘If Bob is happy, then Octo is happy’. We can then formulate the obvious notion of *octopoidal-invariance* and prove *WORLD-OCTOPUS GAP*:

*WORLD-OCTOPUS GAP*: If  $\Delta \models \varphi$  and all of  $\Delta$  are octopus-invariant and satisfiable, then  $\Delta \ \& \ \varphi$  (the conjunction of all the sentences in  $\Delta$  and  $\varphi$ ) is octopus-invariant.

just as we might prove similar things about time-invariance, location-invariance, speaker-invariance, and so on. Call the separated-out things in our model-structures, like privileged locations, times, and speakers, and norms, *indices*. Singer’s

9. Strictly speaking, we need compactness to legitimate this definition, but it’s easy to see how to modify the definition of *WORLD-NORM GAP* to accommodate cases where there is no (finite) conjunction equivalent to  $\Delta \ \& \ \varphi$ .

10. For a classic discussion of this point in an earlier context, and the inspiration for our example, see Pigden (1989) and (2010b).

result is an instance of a general fact about the persistence of index-invariance across implication.

The generality of this fact is important since it means that *WORLD-NORM GAP* only captures Hume's dictum if Singer has appropriately interpreted the descriptive in terms of norm invariance. So, the plausibility of *WORLD-NORM GAP* as an explication of the Hume's dictum depends on the plausibility of Singer's taxonomical distinction between the ethical and the descriptive.

This is a general instance of a problem with interpreting model-theoretic semantics. When we give a *semantics* for various notions, we need also ask which features of these expressions should be built into the semantics as corresponding to the meaning or nature of the expressions themselves. For an example in a different context, suppose our index has a designated speaker slot for interpreting the first-person pronoun. We can then develop a logic of the first-person pronoun *à la* Kaplan (1989). But should we allow any object to serve in the speaker role, like a wheel of cheese, a paramecium, or our favorite restaurant? Should we allow non-existent or dead persons like Socrates, Atatürk, or Morgan le Fey to fill that role? This is by no means clear.

If we do so, we have moved away from giving a logic of the first-person pronoun, in a sense, since our actual notion of the reference of a first person pronoun is plausibly always an (existent) agent capable of representing themselves. Kaplan (1989) argues that what we really want, in order to characterize for the logic of the first-person pronoun, is an account of implication as preservation of truth over *proper points*: pairs of worlds and indices where the object filling the speaker role in the index is of the appropriate type. The take home lesson here is that whether or not an account of indexicals counts as a logic of the first-person pronoun depends on what constraints we build into the background semantics.<sup>11</sup> The analogous lesson for treatments of ethical terms will rear its head throughout our discussion.<sup>12</sup>

Similar problems afflict Singer's characterization of the descriptive as the norm invariant. This only works on certain metasemantic views about the meaning of the ethical expressions and only when we do not impose semantic constraints analogous to restricting 'I' to existent agents. Summarizing our point

---

11. Thanks to an anonymous referee for catching our implicit hat tip to Kaplan's point and stressing that we should make it explicit.

12. Importantly, this is not to say that we cannot give a general semantic account of notions like necessity or indexicals which doesn't build in such connections. We can. However, such an account does not correspond to the intuitive meaning of the natural language expressions we're theorizing about when we do not build in these connections. Rather, such accounts correspond to a more general account of the semantic function of notions like alethic modals *tout court*. The important fact is that moving from this type of abstract model-theoretic account to one which actually represents the meaning of natural language expressions in the appropriate way requires building into the semantics features like the ones under discussion.

briefly, WORLD-NORM GAP captures an interesting form of ethical autonomy only if Singer's taxonomy is plausible as an account of the distinction between the ethical and the descriptive; if it's not, then even though WORLD-NORM GAP will still hold—it's a basic result about index-invariance, after all—it will fail to capture the intuitive gap between the ethical and the descriptive.

### 2.3. *Specific Worries about Ethicality as Norm Invariance: The Descriptive*

In support of our claim that the crux of the matter for accounts like Singer's is their taxonomic distinction between ethical and descriptive claims, we will now show that WORLD-NORM GAP holds even for accounts where Hume's dictum intuitively and paradigmatically fails. WORLD-NORM GAP thus fails to capture how certain metaethical views deny a gap between 'is'-facts and 'ought'-facts.

Suppose, following Searle (1964), that we can impose ethical demands on ourselves by uttering 'I promise to give you 5 bucks' and this follows from facts *analytic* of the notion of promising. If Searle is right, we can *analytically infer* from the existence of a piece of descriptive behavior—an utterance of 'I promise to give you five bucks' (in the appropriate circumstances)—to the existence of an ethical demand on ourselves. Presuming we already restrict the set of worlds in such a way as to respect the analytic connection between promising behavior and promissory obligations, coming to accept the conclusion doesn't restrict the set of possibilities consistent with our beliefs. Nevertheless, the inference adds an explicit ethical proposition that was antecedently only implicit in the combination of putatively analytic facts about the ethical and descriptive facts.<sup>13</sup>

But it is still true, even given Searle's view, that if a set of sentences is norm invariant and satisfiable, then the logical consequences of this set are likewise norm invariant.<sup>14</sup> The key fact here is that promissory behavior, given the analytic connection between it and our obligations, is not norm invariant for Searle. The premises here—facts about speech behavior—thus count as ethical facts on Singer's taxonomy. Intuitively, this is incorrect; these are intuitively descriptive facts. Singer's taxonomy here salvages WORLD-NORM GAP by treating an intuitively descriptive fact as ethical.

We find similar examples in considering conventionalist metasemantic views on which there is generally no substantive gap between 'is'-facts and 'ought'-

---

13. See Pigden (in press) for a useful discussion of Searle's examples and the crucial point that we should treat the connection between promissory obligation and promise-oriented speech behavior as an analytic connection. This has occasionally been missed by responses to Searle's original argument.

14. Of course, it is only with the analytic connections articulated as part of the premises that it logically follows; even without this it analytically follows. Note that, presumably, the facts which underwrite this implication are themselves true at every acceptable world-norm pair and so are themselves norm invariant and thus descriptive on Singer's taxonomy.



facts. Such a view, in its simplest form, holds that the contents of the field of world-norm pairs  $\langle w, n \rangle$  are related in the following way: the norms figuring in  $n$  correspond to the behaviors of the agents in  $w$ .<sup>15</sup> We can call such world-norm pairs *proper* in analogy to Kaplan's restriction on possible occupiers of the agent role. Consider the claim that people tend to avoid stealing and say things like 'stealing is impermissible'. This claim about the behavior of people is not norm invariant over the proper worlds. To see this, consider that a proper pair which contained norms condoning stealing would also—since it is proper—be a world in which people stole and uttered endorsements of stealing. WORLD-NORM GAP holds for this view, just as it does on Searle's view. What potentially fails is the adequacy of Singer's taxonomy since it marks people's tendency to avoid the physical act of stealing as an ethical fact, along with their "moral" pronouncements.

Both these views are ones on which WORLD-NORM GAP should fail. The fact that Singer's view treats them as views with a bizarre understanding of the ethical, but one consistent with WORLD-NORM GAP, shows that it is really the taxonomy doing the work here. That's simply not what these views are. That is, Singer's claim that world-norm gap captures the target gap between 'is'-facts and 'ought'-facts is correct *only if* his definition of ethical facts as *just those which are norm invariant* is plausible and sufficiently neutral between reasonable alternative metasemantic pictures. His claim thus relies on all and only intuitively ethical facts being norm invariant, given plausible and widespread presumptions about the meaning of ethical terms. We think, however, that this claim is far from obvious; we have just given some cases where Singerian taxonomy labels clearly descriptive facts ethical. We will bolster this point by giving examples of cases where Singerian taxonomy labels clearly ethical facts descriptive.

#### 2.4. Specific Worries about Ethicality as Norm Invariance: The Ethical

We start with Terence Cuneo and Russ Shafer-Landau's (2014) thesis that there are *moral fixed points*: substantive moral propositions are also analytic truths.<sup>16</sup> They include examples such as that it is wrong to torture others just because they have inconvenienced you, that it is wrong to sexually assault a child solely to indulge one's lust, and that the interests of others are sometimes morally

15. For details about how to formulate views of this type correctly, see Einheuser (2006) and Woods (2016; in press). Note that when we interpret Gibbard's view literally in terms of plans, some such restriction like that to proper worlds is already called for—not everything can be a complete plan.

16. Strictly speaking Cuneo and Shafer-Landau label these conceptual truths in order to allow that they may have substantive content, as opposed to analytic truths, which they treat as having only formal content. We're a bit more permissive about analyticity and modifying our point to be more restrictive would be trivial, so we'll not worry about the distinction between analytic and conceptual truths.

weightier than our own.<sup>17</sup> If any such proposition is analytic, then it will be included in every complete set of norms and thus be norm invariant. That is, it will be descriptive by Singer's lights.

Mirroring our discussion above, we get easy *intuitive* violations of WORLD-NORM GAP by considering arguments like:

- (a) It is wrong to torture someone just because they have inconvenienced you,
- (b) Your only interaction with Bob so far was one of inconvenience,
- (c) Therefore, it is wrong to torture Bob.

where the conclusion is clearly one intended to be norm-variant in the intuitive sense, yet where all the premises are norm invariant. Singer's account takes a definite stand on these arguments, even given views on which there are moral fixed points. If (a) is a moral fixed point (and hence norm invariant), Singer's taxonomy entails that (c) is also norm invariant. So (c) must be descriptive.

But this conclusion is a *paradigmatically* ethical proposition.<sup>18</sup> We thus need either to rule out accounts of the meaning or metaphysics of the ethical like the conventionalist one above, Searle's, or Cuneo and Shafer-Landau's, or we need to accept that the fact that it is wrong to torture Bob might be descriptive, while the fact that I've promised to give you five dollars is ethical. Both of these options are problematic.

Turning to more common potential conceptual truths, consider the widely-held claim that the ethical supervenes on the descriptive. In one very strong form, this is the claim that it's metaphysically necessary that no two metaphysically possible worlds differ in their ethical component but not their descriptive component. In another strong form, this is the claim that it is conceptually necessary that no two conceptually possible worlds differ in their ethical components, but not their descriptive components. Representing these in Singer's semantics involves banning pairs of world-norm pairs which agree in their worldly component but disagree in their ethical component—which of these we do, and which versions of supervenience are problematic will depend on whether we are taking world-norm pairs to represent metaphysical possibility or an epistemic possibility. Now, assuming supervenience, let *W* be the conjunction of all the worldly

---

17. They also offer "It is *pro tanto* wrong to impose severe burdens on others simply because of their physical appearance" (2014: 405) but we're confident it's ok for you to pay someone burly to move your piano.

18. Of course, the moral fixed points are also paradigmatically ethical! We get similar results when it is not analytically but metaphysically necessary that any system of norms will include some particular ethical proposition and that propositional content is determined by something like sets of metaphysically possible world-norm pairs. So choosing between conceptual and metaphysical construals of world-norm pairs doesn't clearly avoid our worry; either way, there are common and not grossly implausible metaethical views that undermine Singer's taxonomy.

facts about a possible world  $w$ . Let  $N$  be some norm that holds in some pair  $\langle w, n \rangle$  in our set of factual-ethical pairs. Since we've assumed supervenience, there will be no pair  $\langle w, n \rangle$  in our set such that  $N$  is not in  $n$ . Since this means that  $N$  is entailed by  $W$ , just as above, either  $N$  is descriptive or  $W$  is ethical. Neither conclusion is palatable.<sup>19</sup>

Searle's, Cuneo and Shafer-Landau's, and conventionalist accounts are not universally accepted and neither is supervenience in the strong forms just mentioned. Far from it. Our point is just that if any of these were right, then these patently ethical propositions would come out as norm invariant on Singer's account. These views are contentious, but they're hardly crazy. So we have a highly counterintuitive consequence of his taxonomy. Singer's notion of an ethical proposition is not sufficiently neutral between background metasemantic and metaethical principles since, on several non-implausible views, it either treats clearly descriptive claims about behaviors as ethical or treats plausibly analytic ethical claims as descriptive.

### *2.5. The Priority of Ethical Theory*

The truth or falsity of views like Searle's, Cuneo and Shafer-Landau's, and conventionalism cannot be generally assumed either in debates about Hume's dictum or in developing an acceptable taxonomical distinction between the ethical and the descriptive.<sup>20</sup> These matters are more general than the particular views which cause problems for Singer.<sup>21</sup> The upshot is that glossing the descriptive in terms of ethical invariance makes sense only if we allow, as a matter of meaning or metaphysics, (1) norms to vary completely independently of worlds and (2) for no norms to hold at every world-norm pair. Given the contentiousness of both (1) and (2), Singer's claim that we can treat norm invariance as a non-controversial account of the descriptive is simply false; it is a matter of substantive controversy

---

19. We agree with an anonymous referee that this point holds for the stronger forms of supervenience, but not the weaker view that it is a conceptual necessity that no two metaphysically possible worlds differ with respect to their ethical component without differing in their descriptive component. As long as we allow that the conceptually possible world-norm pairs can come apart from the metaphysically possible world-norms pairs (though see below), there will be space for a non-trivial version of supervenience even given Singer's Gibbardian framework. However, we think that the stronger versions of supervenience are common enough that our point is troublesome for Singer.

20. Or, in fact, any of many other views that would serve to illustrate our point.

21. We have no objection to letting an account do a bit of tidying around the edges of a generally acceptable taxonomy or taking a stand on funny cases of the is-ought gap. Our point is that the problematic views and the problems they cause are too widespread and plausible to account for them this way. See, for dramatic demonstration of this point, the discussion of Schurz's result about Restall and Russell's taxonomy below.

whether it is even extensionally adequate as an account of the descriptive.<sup>22</sup>

More generally, the taxonomical distinction between ethical and descriptive propositions involves facts about the meaning of ethical terms. Since model theory is a model of an underlying intuitive account of the meaning of things like ethical expressions, we cannot, without argument, simply set up a formal semantics where norms and worlds vary completely independently and where no claim is part of every set of norms.<sup>23</sup>

Of course, Singer might claim that since his gloss on the is-ought gap is meant to be semantic in the sense of deriving from the meaning of various expressions, it should not make presumptions about the underlying distribution of norms. Perhaps we really ought to assume a semantics where norms and worlds vary freely in order to make sense of the meaning of ethical expressions. Unfortunately, this still fails to do justice to the sense in which meaning can be beholden to background semantical, metasemantical, and metaethical issues. Suppose, for example, that Searle or the conventionalist is correct about the relationships of the ethical component of a world-norm pair to the worldly component. It is then potentially part of the meaning, in the sense of being analytic of, various ethical notions that they co-vary with underlying conventional behaviors of the agents making judgments involving them.<sup>24</sup> Put simply, it is not uncontroversial whether a proper semantics for ethical notions should allow norms and worlds to vary freely, as our discussion above and the earlier discussion of Kaplan's use of proper worlds shows. This is a matter for metasemantical and metaethical dispute.

If we are right here about the taxonomy, then there will be many entailments of intuitively ethical facts from descriptive facts since norm invariance only corresponds with the intuitively descriptive in the absence of background constraints on the field of admissible world-norm pairs. So, Singer's claim that

---

22. Singer could object that his definition can be modified to accommodate this problem, but it is a rather easy problem to generalize. Hume's dictum does not deny that there can be worldly consequences of facts about the existence or non-existence of norms in general or, even, worldly consequences of facts about the existence or non-existence of certain particular norms.

23. See Burgess (2008), Schroeder (2015), and Yalcin (in press) for defense of this point about the role of semantics. It might be thought that Singer's aim is to just characterize Hume's dictum on a fully Gibbardian account of the meaning of ethical terms. But this is not his intention, as his closing remarks about the applicability of his solution to any reasonable semantics and his willingness to extend his definition of norm invariance to any reasonable semantics for moral terms suggests. See the end of Singer's Section V. However, it is possible to take his view as an explication of what the Hume's dictum would come to if Gibbard's permissiveness about norms were allowed. Though, even then, there is a substantive worry about whether taking seriously the planning interpretation of norms requires restricting the field of world-norm pairs.

24. This does not mean that it is analytic of such notions that they co-vary with the underlying conventional behaviors at a context which is being evaluated. This would confuse conventionalism with naive subjectivism.

world-norm gap captures *the* sense in which Hume thought we can't get 'is'-facts from 'ought'-facts fails. The formal notion of norm invariance which is persistent over implication only captures the descriptive when we ban common analytic or metaphysical constraints on possible sets of world-norm pairs. We now turn to discussing Restall and Russell, the other main advocates of this style of approach.

### 3. Restall and Russell's Account

#### 3.1. Restall and Russell's Approach: ETHICAL FRAGILITY

Russell and Restall attempt a general account of a range of logical 'barrier' theses, including a barrier between natural and ethical sentences, as well as barriers between particular sentences and general sentences, past and future, contingent and necessary. Their approaches to each of these barriers is different, since they too are aiming for a substantive account that attends to the meanings of the terms on either side of their barriers. Their account of the ethical barrier thesis we are concerned with starts with the property of *fragility*.

A sentence  $\varphi$  is *R-fragile*—*R* indicating some accessibility relation between models—if and only if for every model  $\mathfrak{M}$  in which it is true, there is some *R*-accessible model  $\mathfrak{M}^*$  where  $\varphi$  is false. Universal sentences are *extension-fragile* because they can be falsified by extensions to the model, for instance by adding a white raven to a domain full of black ones. Necessary sentences are also *extension-fragile*, in a possible worlds semantics, since they can be falsified by adding additional worlds. A sentence that is not *R-fragile*—an *R-stable* sentence—does not imply any sentence that is *R-fragile*. The proof is easy. Let  $\varphi$  be *R-stable* and  $\psi$  be *R-fragile*. *R-stable* sentences are consistent by definition. If  $\varphi$  and  $\psi$  are not jointly inconsistent, then any model of  $\varphi$  witnesses the failure of the implication of  $\psi$  from  $\varphi$ . If they're jointly consistent, we have a model  $\mathfrak{M}$  of both  $\varphi$  and  $\psi$ . Since  $\psi$  is *R-fragile*, we can extend the model to some  $\mathfrak{M}^*$  where  $\psi$  is false. Since *R-stable* sentences true in  $\mathfrak{M}$  are true in  $\mathfrak{M}^*$ ,  $\varphi$  is true in  $\mathfrak{M}^*$  and we have our counterexample.

Russell and Restall define an ethical sentence as one that is fragile in one of the following two ways: for every model  $\mathfrak{M}$  in which  $\varphi$  is true, either there is some model  $\mathfrak{M}^*$  in which  $\varphi$  is false that can be derived from  $\mathfrak{M}$  by adding (natural) worlds to the model (*fragility<sub>N</sub>*) and extending the relevant accessibility relation for normative expressions or there is some model  $\mathfrak{M}^*$  in which  $\varphi$  is false which can be obtained from  $\mathfrak{M}$  by changing which worlds are ideal (rejiggering the relevant accessibility relation for normative expressions) in the sense familiar from deontic logic (*fragility<sub>I</sub>*). They define *descriptive* sentences as those not ethically fragile in either sense. This means that for every model  $\mathfrak{M}$  in which a descriptive sentence  $\varphi$  is true, it will remain true in every model  $\mathfrak{M}^*$  that is ob-

tained either choosing new ideal worlds from among the worlds in  $\mathfrak{M}^*$  or adding new worlds to  $\mathfrak{M}$  and making corresponding changes to normative accessibility. The proof that descriptive sentences do not validly entail ethical sentence then goes through as above.

The informal gloss they offer of fragility<sub>N</sub> helps to get an idea of what they're up to:

... that ethical sentences are fragile over changes in situation through which descriptive sentences are preserved is plausible. Consider a very simple situation in which the only thing that happens is that Alice intentionally hits (and hurts) Bob. This is prima facie a situation in which Alice does something wrong. But now 'extend' (in an informal sense) the situation to one in which Alice and Bob are in training for a boxing tournament, and Bob demands (in the name of good training) that his partners try as hard as they can to win. Suddenly it is plausible that what Alice did was not wrong after all, though the description we gave of the first situation is still a true one. Now imagine a third situation, an extension of both the first and second ones. This time we add Candy, a suicidal anti-boxing protester who has informed Alice that she will kill all three of them if Alice hits Bob. Surely Alice would be wrong to hit Bob in that situation. And yet the description we gave of the second situation is still true of this third one. So we think it plausible that the sentence 'it is wrong for Alice to hit Bob' is fragile over these changes in situation, though (for example) 'Alice hits Bob' is not.<sup>25</sup> (Restall and Russell 2010: 256)

Since their account of ethical fragility is disjunctive, we need, for a counterexample to their taxonomy, examples of intuitively ethical claims that violate both disjuncts. We need examples true in a model  $\mathfrak{M}$  which cannot be falsified either by choosing a different set of ideal worlds (violations of fragility<sub>I</sub>) or by adding new worlds to  $\mathfrak{M}$  (violations of fragility<sub>N</sub>). We will construct such an example, starting with a general way of violating fragility<sub>N</sub>.

---

25. Russell and Restall have switched here from 'ought' and 'permission' talk to talk about wrongness. But this brings additional complications unless they assume the contentious hypothesis that to do something wrong is to do what you morally ought not. Can they do this? This depends on their interpretation of their models, on which see the discussion of fixed points below.

### 3.2. Problems for Fragility<sub>N</sub>

The basic problem with fragility<sub>N</sub> is that not all intuitively ethical properties are fragile in this sense.<sup>26</sup> That we ought to do something and that we are permitted to do something (along with properties like being morally wrong, being irrational, etc.) are all-things-considered properties of actions, and as such they are highly sensitive to changes in descriptive features of the situation in which they obtain. Add one demanding madman with an explosive and the ought facts change pretty quickly.

It is less plausible that other ethical properties are fragile<sub>N</sub>, such as intrinsic goodness, badness, or wrongness. Suppose that experiences of pleasure are intrinsically good, and that the torturing of children is intrinsically *pro tanto* wrong. Then the fact that Alice is experiencing pleasure will be intrinsically good, and it will remain intrinsically good whatever else is going on in her situation. Bob's torturing of the child will be intrinsically *pro tanto* wrong, and it will remain so whatever else is going on in his situation. In this respect 'Alice's experience of pleasure is intrinsically good' and 'Bob's torturing of the child is intrinsically *pro tanto* wrong' will be non-fragile<sub>N</sub> across the same range of *descriptive extensions* as 'Alice hits Bob.' In fact, it will be less fragile<sub>N</sub> than a descriptive *extrinsic* sentence. Suppose Alice is actually Lucy's girlfriend. Being Lucy's girlfriend is an extrinsic fact about Alice. Consider 'Lucy's girlfriend hit Bob, therefore Lucy's action was intrinsically wrong.' In this case the antecedent but not the consequent is fragile<sub>N</sub>.

Since the Humean intuition is one about ethical properties in general, not merely ought and permissibility facts, this points to a serious problem. If true, Hume's dictum presumably applies to 'local' ethical facts just as much as it does to 'all things considered' ethical facts. This gives us good reason to think that this first kind of fragility, fragility<sub>N</sub>, cross-cuts the property of being ethical. That is, there is a problem with modeling the notion of an ethical sentence by means of fragility<sub>N</sub>. Unless all of ethics is exhausted by oughts and permissions—and this is unlikely to be true—then their taxonomy is inadequate on its own as a characterization. This is a structural point, not a point about particular metaethical views like those discussed above. It doesn't rely at all on first-order analytic claims like Searle's or Cuneo and Shafer-Landau's.

But Russell and Restall's characterization of ethical fragility is disjunctive, so for a full taxonomical counterexample we need an example which also violates the second sense of fragility: fragility across changes in which worlds are relatively ideal. Change which worlds are ideal, and you change whether it is true that it is permissible for Alice to hit Bob. You might alternatively change which

---

26. Vranas (2010) presents analogous doubts about whether all natural sentences are 'descriptive' in the associated sense.

actions are intrinsically wrong, and which states of affairs are intrinsically valuable. Fragility<sub>I</sub> is not vulnerable to the previous concerns, so our counterexample is incomplete.

### 3.3. Problems for Fragility<sub>I</sub> and a Full Taxonomic Counterexample

Fragility<sub>I</sub> is, however, vulnerable to the sort of worries we pressed against Singer above. For example, consider again Cuneo and Shafer-Landau's (2014) view that there are some 'moral fixed points': propositions such that it is conceptually necessary that if any ethical sentence is true, these sentences are true. For a historically fun example, we might think that the fact that it is morally wrong to destroy the world in order to avoid scratching your little finger is a moral fixed point. It would follow that if there are any ethically ideal worlds, in them no one would destroy the world in order to avoid scratching their little finger.

If one of these moral fixed points involves local intrinsic ethical properties, and some plausibly do, then it will fail to be fragile in both respects. It will be impervious to extensions of the model (fragility<sub>N</sub>) and to changes in which relative worlds are ideal (fragility<sub>I</sub>). For example, this might be the case pain: pain being intrinsically disvaluable seems perhaps a conceptual truth and thus one which is not fragile<sub>I</sub> on an adequate interpretation of our ethical concepts. So we have a counterexample to Russell and Restall's taxonomical thesis. Examples of intuitively ethical claims like 'pain is disvaluable', which involve local intrinsic properties and are possibly conceptual truths, show that that even the disjunctive notion of fragility doesn't yield a plausible account of the ethical/descriptive divide.

A related worry is raised in Schurz (2010b). Suppose it is (conceptually) necessary that if we must do something, then we ought to do that thing. Call this MUST-UGHT. In its more obvious form, MUST-UGHT claims that if we are permitted to do something, then it is possible to do that thing. MUST-UGHT is true at a world  $w$  in a model if the worlds which are ideal from  $w$  are a subset of the worlds that are possible from  $w$ . Let's suppose that MUST-UGHT is an analytic truth; it certainly is potentially such. Then we can restrict our attention to models which validate it. That is, proper models where the ideal worlds are subsets of the possible worlds.

Now, let  $\varphi$  be a descriptive sentence. Then ' $\varphi$  is necessary' is likewise descriptive. Is ' $\varphi$  is obligatory' ethical? Intuitively, yes. So it should satisfy Russell and Restall's definition of an ethical truth. But it does not: to see this, note that there will be perfectly good models in which  $\varphi$  is a necessary truth. In such a model, ' $\varphi$  is obligatory' is not fragile<sub>N</sub>. Let  $w$  be a world in which ' $\varphi$  is obligatory' is true. Suppose we added worlds, extending perhaps the normative accessibility relation, and thereby made ' $\varphi$  is obligatory' false at  $w$ . This means we have added worlds in such a way that at some ideal world  $v$  (from the perspective of



$w$ ),  $\varphi$  is false. By MUST-UGHT,  $v$  must be possible with respect to  $w$ . But then the failure of  $\varphi$  at  $v$  witnesses the failure of ' $\varphi$  is necessary' at  $w$ . That, however, would mean that ' $\varphi$  is necessary' is ethical, not descriptive. Which, of course, it ain't.

It is also not fragile<sub>I</sub>. Suppose we choose a new set of ideal worlds, leaving the rest of the model alone. Since our new model will also validate MUST-UGHT, from the point of view at any world, the ideal worlds will be a subset of the accessible worlds. Since ' $\varphi$  is necessary' is descriptive and true at  $w$ , it will be true throughout the new set of ideal worlds, and hence ' $\varphi$  is obligatory' will be true at  $w$  as well. Since ' $\varphi$  is obligatory' is not ethically fragile, it is descriptive. And this point holds for *any claim* of the form ' $\varphi$  is obligatory' where  $\varphi$  is descriptive. This, however, is a very strange result. So, in order for Russell and Restall's characterization of the distinction between the ethical and the descriptive to work out, principles like MUST-UGHT cannot be analytic truths. This, however, seems to be a very strong commitment for a supposedly neutral characterization of the is-ought gap.

### 3.4. *The Priority of Ethical Theory*

So both our examples and Schurz's result point to the fact that the plausibility of the formal characterization of ethicality is dependent on the existence, or not, of background conceptual truths. Russell and Restall could claim, in response to these cases, that they are only concerned with the *logical* relationship between descriptive and ethical facts, as represented in their model theory. This is consistent with denying even conceptual truths about which sentences are ethical if any are, and hence with denying the principle that there is some sentence  $\varphi$  such that if any ethical sentence is true,  $\varphi$  is true. The fixed point sentences are entirely optional in the model theory, as is MUST-UGHT. Consequently, there is a model  $\mathfrak{M}$  in which it is morally permissible to destroy the world to avoid scratching your little finger, and so there will be no descriptive sentence that entails any ethical sentence relevantly related to  $\mathfrak{M}$ . Similarly with putatively analytic principles like MUST-UGHT.

But this reply just points to a shortcoming in Russell and Restall's autonomy thesis. It is consistent not just with the Cuneo and Shafer-Landau moral fixed points hypothesis, but with the hypothesis that it is conceptually necessary that, if any ethical sentence is true, then precisely one fully determinate ethical theory is true—and with the extreme metaphysical hypothesis that it is in the nature of ethical properties to be such that all ethical properties are identical to some natural property. But it is theses like these—as we suggested above—that philosophers often take themselves to be disputing when they are arguing about autonomy. Plausibly enough, our taxonomical line between the ethical and the descriptive shouldn't be drawn differently depending on which of them is true.

It also points to a significant worry. Their intuitive gloss on their thesis, quoted above, suggests that certain conceptual relations among ethical terms are built into their interpretation of their models. For example, the connection of wrongness with ought facts and permission facts needs to be a conceptual truth if their switch from one locution to the other in the gloss is permissible. However, once some conceptual connections between ethical terms are presumed, we can legitimately ask why others, such as *MUST-UGHT* or the intrinsic disvalue of pain, are ignored. Even the interdefinability of ought and permissibility is a non-trivial thesis about the meaning of 'ought' and 'permitted', after all.

Essentially, what Russell and Restall have shown is that *if* all and only ethical propositions meet their structural conditions, and all and only descriptive propositions meet their structural conditions, then descriptive propositions don't entail ethical propositions. What is needed is an argument that the minimal structure Russell and Restall presume about ethicality and descriptivity really capture essential features of these types of propositions. But, as we noted above, this is a non-trivial metaethical claim about ethicality that we have yet to find an argument for. Given this fact, their thesis is *not* metaethically neutral in the sense described above.

Logical characterizations of ethical autonomy, as the weakest types of characterization, underdetermine some of the most interesting metaethical questions that have been thought to concern autonomy. For instance, metaphysical questions about the relations of grounding, reduction, or identity between descriptive and ethical facts are unaddressed. Epistemic questions about the immunity of ethical hypotheses from descriptive evidence are also left unaddressed. Many of the philosophical disputes which turn in one way or another on some version of the autonomy of ethics would not be settled even by a logical characterization of the thesis which received universal acceptance—and certainly not if that logical characterization presupposed the truth of the metaphysical or conceptual or epistemic claim at issue.

#### 4. Lessons Learned

Part of what explicitly motivates Singer to give his account, and presumably part of what motivated Russell and Restall to construct theirs, is the prospect of an account of the gap between *is* and *ought* which also permits the development of an illuminating deontic logic—unlike the earlier account given by Pigden (1989).<sup>27</sup> We will not weigh in on this motivation (as criticism or desiderata), except to note that deontic logic, like tense logic or modal logic, is really a formal theory

---

27. Though Schurz (1997) develops the Pigden approach in such a way as to permit illuminating deontic logics, mainly by treating 'ought' as a logical operator and modifying Pigden's vacuousness criterion. See also Schurz (2010a) and Pigden (2010b) for discussion.

of the meaning of ethical and ethical expressions. And as such, the semantic base over which the logic is defined needs to accommodate the analytic features of the notion being theorized.<sup>28</sup>

For example, when we develop a logic of alethic modals, we restrict the accessibility relation between worlds so that every world “sees” itself—that is, we require that the accessibility relation on possible worlds be reflexive. We then can and do argue about whether it should also be transitive, Euclidean, etc. Some of these arguments take place at the level of claims about the meaning of alethic modals, corresponding to claims about the meaning of ‘necessary’ itself; some take place at the level of fundamental metaphysics, corresponding to claims about the domain of application we want to apply ‘necessary’ over.<sup>29</sup> Constraints motivated by the first type of argument, such as the reflexive property, correspond to features of the meaning of ‘necessity’ itself. A semantics for a notion like necessity must take into account these analytic connections. It need not build metaphysical connections into the semantics itself, but if it proceeds by using something like metaphysically possible worlds as part of the semantics, then such features will also turn up, implicitly, in the semantics.

When giving a theory of ethical expressions, one introduces operators or connectives and specifies some rules and axioms for them. It is then an open question to what extent the new operators or connectives and their rules correspond to the properties one is “giving a logic of.” In the case at hand, it is an open question how well extremely permissive formal accounts of deontic or ethical possibility do in capturing our intuitive notion of such. What Russell and Restall are doing, for example, is providing a purely model-theoretic characterization of ‘ought’ and ‘permitted’. Then they prove a ‘barrier thesis’ in their new formal system. Their hope is that the formal property of ethical fragility corresponds closely enough to the intuitive property of being ethical for their barrier thesis to have philosophical significance. Presumably, though they are not as explicit about this as Singer, their account should not take on substantive meta-ethical commitments like those we have mentioned above; it should be above the particular fight and fray of metaethical dispute.

This strategy would work, though, only if the distinction between being an ethical and a descriptive proposition admits of an illuminating formal characterization independently of these debatable claims. This is a plausible ambition for the relatively formal distinction between particular and universal, or perhaps between contingent and necessary. But it seems to us that there are simply too few ethically uncontentious hypotheses about ethical structure itself for any model-

---

28. Arguably, this is true even for purely logical expressions like ‘and’ and ‘not’, but the case for this would distract here. It’s clearly true for cases like those we’re discussing.

29. Compare, for example, the arguments against the S4 modal axiom (the transitivity axiom) by Salmon (1989) to those by Dummett (1993).

theoretic account of the ethical/descriptive distinction to remain sufficiently neutral.<sup>30</sup> The formal distinctions given by Russell and Restall, and by Singer, should cross-cut central questions about ethical taxonomy. But they don't.

What we have claimed and demonstrated repeatedly is that the intuitive adequacy of the formal distinction between the ethical and descriptive, on either Singer's characterization or Russell and Restall's, itself is beholden to a background view about what the conceptual truths about the meaning of ethical terminology. In particular, what our above examples show is that on many views of the conceptual truths governing the ethical, these formal characterizations mischaracterize the distinction between the ethical and the descriptive. This happens on views like Cuneo and Schafer-Landau's where there are moral fixed points, views which accept plausible conceptual claims about the ethical like *MUST-UGHT*, and views like Searle's and the conventionalist's where natural and ethical facts are analytically connected. When ethical possibility is so constrained by such conceptual truths, the 'descriptive' becomes surprisingly wide, including all kinds of paradigmatically ethical claims like those discussed above.

This is not to say that we can't draw the formal distinctions Singer or Russell and Restall have in mind even if there are some substantive conceptual constraints on the ethical. We can. However, the formal distinction doesn't represent the intuitive distinction between the ethical and the descriptive in the presence of certain background constraints, many of which are themselves plausibly part of our concept of the ethical.

We conclude, then, that it is far from clear that the distinction between the ethical and the descriptive admits of an illuminating formal characterization and it is clear that the formal characterizations given by Singer and by Russell and Restall are not metaethically neutral. They succeed at demonstrating that the ethical is substantively insulated from the descriptive if ethical possibility is unconstrained by the sort of conceptual truths given above, but only if it is so unconstrained. It would be better to develop an account of the autonomy of ethics where Hume's dictum is a substantive feature of our (meta)ethical theory, such as the metaphysical characterization developed by Maguire (2015) and refined by Woods (2017).

Finally, to ward off an objection, remember that we are not here interested in the possibility of characterizing a purely logical gap between the ethical and the descriptive. Such a gap could reasonably allow any selection of norms or any set of ethically ideal worlds, since neither 'ought' nor 'permitted' nor 'wrong' are plausibly logical terms and so we do not hold their substantive meaning fixed when we investigate what entails what.<sup>31</sup> As we suggested above, they succeed

---

30. Again we echo Sayre-McCord (1986).

31. We presume here the commonly accepted Tarskian account of logical consequence in terms of preservation under systematic reinterpretation of the non-logical expressions in a

at this, but this sort of gap is rather uninteresting, holding for many distinct semantical categories. It is *only if* the essential features of the concept of the ethical can be cashed out in a neutral fashion in formal terms that the sort of formal picture given by Russell and Restall, and by Singer, would be illuminating in the way that many, including ourselves, hope that Hume's dictum is.

## Acknowledgements

Thanks to a couple of helpful referees, Charles Pigden, Gillian Russell, Geoff Sayre-McCord, Paolo Santorio, and Daniel Singer for discussion.

## References

- Brown, Campbell (2014). Minding the Is-Ought Gap. *Journal of Philosophical Logic*, 43(1), 53–69.
- Burgess, John (2008). Tarski's Tort. In *Mathematics, Models, and Modality* (149–168). Cambridge University Press.
- Cuneo, Terence and Russ Shafer-Landau (2014). The Moral Fixed Points: New Directions for Moral Nonnaturalism. *Philosophical Studies*, 171(3), 399–443.
- Dummett, Michael (1993). Could There Be Unicorns? In *The Seas of Language* (328–348). Clarendon Press.
- Einheuser, Iris (2006). Counterconventional Conditionals. *Philosophical Studies*, 127(3), 459–482.
- Gibbard, Allan (1990). *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Harvard University Press.
- Kaplan, David (1989). Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein (Eds.), *Themes From Kaplan* (481–563). Oxford University Press.
- Karmo, Toomas (1988). Some Valid (but No Sound) Arguments Trivially Span the 'Is'-'Ought' Gap. *Mind*, 97(386), 252–257.
- Maguire, Barry (2015). Grounding the Autonomy of Ethics. In Russ Shafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 10, 188–215). Oxford University Press
- McPherson, Tristram (2008). Metaethics & the Autonomy of Morality. *Philosophers' Imprint*, 8(6), 1–16.
- Pigden, Charles R. (1989). Logic and the Autonomy of Ethics. *Australasian Journal of Philosophy*, 67(2), 127–151.
- Pigden, Charles R. (Ed.) (2010a). *Hume on Is and Ought*. Palgrave Macmillan
- Pigden, Charles R. (2010b). On the Triviality of Hume's Law: A Reply to Gerhard Schurz. In Charles R. Pigden (Ed.), *Hume on Is and Ought* (217–238). Palgrave

---

putative entailment.

Macmillan.

- Pigden, Charles R. (in press). Hume On Is and Ought: Logic, Promises and the Duke of Wellington. In Paul Russell (Ed.), *Oxford Handbook on David Hume*. Oxford University Press.
- Prior, Arthur Norman (1960). The Autonomy of Ethics. *Australasian Journal of Philosophy*, 38(3), 199–206.
- Russell, Gillian and Greg Restall (2010). Barriers to Implication. In Charles R. Pigden (Ed.), *Hume on Is and Ought* (243–259). Palgrave MacMillan.
- Salmon, Nathan (1989). The Logic of What Might Have Been. *Philosophical Review*, 98(1), 3–34.
- Sayre-McCord, Geoff (1986). Deontic Logic and the Priority of Moral Theory. *Noûs*, 20(2), 179–197.
- Schroeder, Mark (2016). Is Semantics Formal? In *Expressing Our Attitudes* (209–224).
- Schurz, Gerhard (1997). *The Is-Ought problem: An Investigation in Philosophical Logic*. Springer.
- Schurz, Gerhard (2010a). Comments on Pigden. In Charles R. Pigden (Ed.) *Hume on Is and Ought* (260–267). Palgrave-Macmillan.
- Schurz, Gerhard (2010b). Comments on Restall, Russell, and Vranas. In Charles R. Pigden (Ed.), *Hume on Is and Ought* (268–272). Palgrave Macmillan.
- Searle, John (1964). How to Derive "Ought" From "Is". *Philosophical Review*, 73(1), 43–58.
- Singer, Daniel (2015). Mind the Is-Ought Gap. *Journal of Philosophy*, 112(4), 193–210.
- Vranas, Peter (2010). Comments on 'Barrier to Implication'. In Charles R. Pigden (Ed.) *Hume on Is and Ought* (198–217). Palgrave-Macmillan.
- Woods, Jack (2016). The Normative Force of Promising. In Mark C. Timmons (Ed.), *Oxford Studies in Normative Ethics* (Vol. 6, 77–101). Oxford University Press.
- Woods, Jack (in press). The Authority of Formality. In Russ Schafer-Landau (Ed.), *Oxford Studies in Metaethics* (Vol. 13).
- Woods, Jack (2017). On Vacuous Grounding: The Case Study of Ethical Autonomy. Manuscript in preparation.
- Yalcin, Seth (in press). Semantics as Model-Based Science. In *The Science of Meaning*. Oxford University Press.