

## Non-Naturalism and Reference

Suikkanen, Jussi

*License:*

Creative Commons: Attribution-NonCommercial-NoDerivs (CC BY-NC-ND)

*Document Version*

Peer reviewed version

*Citation for published version (Harvard):*

Suikkanen, J 2017, 'Non-Naturalism and Reference', *Journal of Ethics & Social Philosophy*.  
<[http://www.jesp.org/PDF/Non-Naturalism%20and%20Reference\\_final.pdf](http://www.jesp.org/PDF/Non-Naturalism%20and%20Reference_final.pdf)>

[Link to publication on Research at Birmingham portal](#)

**Publisher Rights Statement:**

The Journal of Ethics and Social Philosophy is founded on the principle of publisher-funded open access. Public access to articles in the Journal of Ethics & Social Philosophy is free of charge, at no charge to authors, and are available to all readers under the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 license.

**General rights**

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

**Take down policy**

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# 1 Non-Naturalism and Reference

*Jussi Suikkanen*

Final author copy; to be published in *Journal of Ethics & Social Philosophy* (www.jesp.org).

## 1. Introduction

According to metaethical realists, some actions, character-traits, states of affairs, and the like (hereafter simply ‘objects’) have stance-independent normative properties: there are actions we ought to do and outcomes that are good in a way that does not depend on our attitudes towards them. Metaethical realists then disagree about the nature of these properties. The naturalists think that they are ordinary natural properties: causally efficacious, *a posteriori* knowable, and usable in the best explanations of natural and social sciences.<sup>1</sup> The non-naturalists, in contrast, argue that they are *sui generis*: causally inert, *a priori* knowable, not a part of the subject matter of sciences, and yet a ‘part of the fundamental nature of the universe’ (Ross 2002 [1930], 29–30).<sup>2</sup>

---

<sup>1</sup> There is a disagreement about the distinguishing features of natural properties. Causal efficacy was emphasized by Lewis (1983: sec. 2), empirical knowability by Copp (2003), and being used in scientific explanations by Little (1994). Moore (1903: 40) and Shafer-Landau (2006: 211) believe that naturalness is a matter of being part of the subject matter of sciences. I will also count disjunctive properties the disjuncts of which are natural properties as natural properties even if they might not satisfy the previous criteria themselves. This is because such properties are not different kind of properties in the way that the non-naturalists assume normative properties are.

<sup>2</sup> Some non-naturalists do not think that normative properties exist metaphysically (see, for example, Parfit (2011, vol. 2, 479), Scanlon (2014, ch. 2), and Nagel (1996, 205)). For my objections to these views, see Suikkanen (2016). The metaphysically robust non-naturalists discussed here include at least Enoch (2011), Huemer (2005), Ross (2002 [1930]), Shafer-Landau (2003), and Wielenberg (2014).

The metaphysical and epistemological objections to non-naturalism have already been discussed extensively.<sup>3</sup> This article focuses therefore on a less discussed metasemantic question.<sup>4</sup> Our natural language contains a number of normative predicates such as ‘ought’, ‘...is a reason to...’, ‘good’, ‘better’, ‘right’, and so on.<sup>5</sup> According to the non-naturalists, there are non-natural normative properties: the property of being what one ought to do, the relation of reasonhood, the property of goodness, and so on. The non-naturalists then believe that the normative predicates refer to the previous normative properties: ‘ought’ refers to being what one ought to do, ‘...is a reason to...’ to the reasonhood relation, ‘good’ to goodness, and so on.

The metasemantic challenge for the non-naturalists is to provide an explanation of how the previous normative predicates manage to refer to the relevant non-natural normative properties. What is the ‘semantic glue’ that binds the relevant predicates to the corresponding properties?<sup>6</sup> There are, furthermore, three adequacy conditions for the solutions to the previous challenge – three further challenges that the non-naturalists must be able to meet. They need to explain, first of all, how the normative words manage to refer to the non-natural properties instead of natural ones. This ‘Contrast

---

<sup>3</sup> See Mackie (1977, ch. 1). For a recent version of the epistemological objection, see Bedke (2009), and for a corresponding metaphysical objection, see McPherson (2012). For responses, see Enoch (2011, chs. 6–7), Huemer (2005, chs. 5 and 8), Shafer-Landau (2003, chs. 3–4 and 10–12), Wedgwood (2007, 6–11), and Wielenberg (2014, chs. 1 and 3).

<sup>4</sup> See Wedgwood (2007: sec. 1.2) and Setiya (2011: 1283). The question of what the referents of normative predicates are is a semantic question, whereas how they came to have those referents is a metasemantic question.

<sup>5</sup> Following Wedgwood (2007: sec. 1.3), I take the defining feature of normative predicates to be that, when we use them to make judgments, there is a requirement of rationality to have certain motivational states. This is why some evaluative concepts such as ‘good’ and deontic concepts such as ‘wrong’ are also normative concepts. I say more about the sense of normative concepts in §5.

<sup>6</sup> The normative concepts need not pick out the same property on each occasion of use. Following Kripke (1979a), we can distinguish between what the concept itself refers to and what a speaker refers to when using the expression. The metasemantic challenge investigated here concerns the former question.

Challenge' has become more pressing recently because it has become evident that, for every non-natural normative property that the non-naturalists posit, there is a necessarily co-instantiated natural property.<sup>7</sup>

Any good object, for example, will have the conjunctive natural property constituted by its all natural properties (it has the property of being F, G, not-H, ..., and Z – it is D<sub>1</sub> for short). Every possible good object will then have a similar comprehensive conjunctive natural property. This means that all possible good objects have the disjunctive natural property of being D<sub>1</sub>, D<sub>2</sub>, ..., or D<sub>n</sub>. And given supervenience, no object that is not good has that property.<sup>8</sup> This is why goodness and being D<sub>1</sub>, D<sub>2</sub>, ..., or D<sub>n</sub> are necessarily co-instantiated. A central part of the Contrast Challenge is then to explain how the predicate 'good' manages to refer to the non-natural property of goodness instead of the underlying necessarily co-instantiated property of being D<sub>1</sub>, D<sub>2</sub>, ..., or D<sub>n</sub>.

The non-naturalists also need to explain why 'good', for example, refers to the property of goodness rather to some other non-natural property such as rightness. Call this the 'Correct Property Challenge'. This challenge is surprisingly pressing concern for two reasons. Firstly, some non-naturalists believe that there are thick normative predicates such as the aesthetic predicates '... has grace' and '... is delicate' and they also assume that these predicates refer to the non-natural properties of grace and delicacy (see Crisp 2005: 82). Here the task for

---

<sup>7</sup> See Kim (1984), Jackson (1998: 122–123), and Gibbard (2003: 94–102). Majors (2005) denies that the base property would be a natural property. As explained in footnote 1, I count all properties formed by Boolean operations from the paradigmatic natural properties to be natural properties too. Jackson (1998: 123) goes onto argue that necessarily co-instantiated properties are identical. If Jackson's criterion of property identity is correct, then the question of how normative predicates could refer to non-natural properties would not arise – there would not be such properties. For a discussion, see Suikkanen (2010).

<sup>8</sup> Supervenience is roughly the thesis that two objects cannot have different normative properties without having different natural properties. Jackson (1998, 119) relied on global supervenience, but Williamson (2001) showed that Jackson's argument requires the truth of strong supervenience.

the non-naturalists is to explain just how the predicate ‘... has grace’ manages to refer to grace rather than delicacy given that the ways in which these predicates are used are almost identical. Many of the same objects are both graceful and delicate and the reactions these properties call for are also very similar. The second reason this challenge is rather pressing is that, as explained in §5 below, it turns out to be the most difficult to respond to.

Finally, the non-naturalists also face a version of the grue problem – call it the ‘Gruesome Normative Property Challenge’.<sup>9</sup> Grueness is stipulated to be the property of being green when examined before certain time *t* in the distant future or being blue otherwise. In the case of colours, one metasemantic challenge is to explain how ‘green’ manages to refer to greenness rather than to grueness. Similarly, the non-naturalists need to explain how the normative predicates manage to refer to the standard normative properties rather than to their gruesome counterparts. Why does ‘good’, for example, refer to goodness rather than to the property of being good if inside this light cone and being cruel elsewhere?

If the non-naturalists cannot respond to the previous challenges, then synthetic versions of naturalism might seem to be more plausible views in metaethics.<sup>10</sup> They could, after all, attempt to explain how normative predicates refer to certain natural properties by relying on the essential qualities of the natural properties such as their causal powers. The synthetic naturalists do not face the Contrast Challenge, and they could try to use the causal connections between normative predicates and natural properties to meet the other two challenges. They could, for example, argue that ‘good’ refers to goodness – a certain natural property – rather than to any other natural property because that property governs causally

---

<sup>9</sup> The grue-paradox, i.e. the ‘new riddle of induction’, was originally formulated by Goodman (1983: 72–81).

<sup>10</sup> According to synthetic naturalists, the sense aspect of the meaning of normative predicates (what cognitive significance they have for the speakers) does not determine their reference. See, e.g., Boyd (1988), Brink (1989), Copp (2000), and Schroeder (2007). Analytic naturalists, in contrast, take the meaning of normative predicates to fix their reference (see §2 below).

the core uses of the term in an epistemically beneficial way (Boyd 1988: 195).

This article attempts to show that the previous assumptions are mistaken. It argues that (i) both non-naturalists and synthetic naturalists have to rely on the same *non-causal* mechanism to explain reference and that (ii) both have *equally good resources* for making use of that mechanism. As a consequence, my conclusion will be that there is no reason to prefer synthetic naturalism to non-naturalism at least on metasemantic grounds.

My argument proceeds in four sections. §2 explains why the non-naturalists cannot rely on basic descriptivist or Fregean metasemantic accounts as a response to the previous challenges. §3 then argues that Ralph Wedgwood's (2001 and 2007, ch. 4) conceptual role semantics will not help the non-naturalists either.

In §4, I argue that the synthetic naturalists cannot give purely causal accounts of how normative predicates refer because of the *qua-problem*, which has not yet received sufficient attention in metaethics. Finally, §5 suggests both (i) that the most promising way for the synthetic naturalists to avoid the qua-problem relies on the *non-causal mechanism of reference magnetism* and (ii) that the non-naturalists too can rely on it to meet the metasemantic challenges they face.

## 2. Descriptivism and Fregean Views of Predicates

Let us begin from why the non-naturalists cannot rely on simple versions of descriptivism. According to them, the meaning of a predicate '*... is P*' consists of a set of descriptions which an object must satisfy in order for it to count as P.<sup>11</sup> These are the descriptions which a speaker must know for being able to use the predicate in question with competence. For example, the meaning of '*... is a bachelor*' is said to consist of the descriptions '*... is male*' and '*... is unmarried*', which a person must satisfy in

---

<sup>11</sup> Descriptivism can be traced back to Frege (1997a [1892]). It has been defended in different forms by Russell (1905), Strawson (1950), Searle (1958), and others.

order to be a bachelor. The idea is then that predicates refer to the property in virtue of which objects satisfy the meaning-constituting descriptions. Thus, because these descriptions pick out in the previous example the property of being an unmarried male, the predicate ‘... is a bachelor’ refers to that property. After all, a person counts as a bachelor in virtue of having it.

The meaning of the normative predicate ‘...is good’ could also be suggested to consist of some descriptions ‘R’ and ‘S’, which a competent speaker must know and which an object must satisfy in order to count as good. The claim would then be that ‘good’ refers to Rness and Sness because objects count as good in virtue of satisfying the descriptions ‘R’ and ‘S’.

This view, sadly, is not available for the non-naturalists. Either the meaning-constituting descriptions are formulated in wholly non-normative vocabulary or they include some normative predicates. If the descriptions are wholly non-normative, then the view entails that normative predicates refer to natural properties.<sup>12</sup> If they include normative predicates, we face the same question: how did these other normative predicates acquire their reference? If the non-naturalists insist that they too acquired their reference via some meaning-constituting descriptions, no progress has been made. Again, either the

---

<sup>12</sup> The resulting view is analytic naturalism. G.E. Moore (1903: §39) famously attributed this view to John Stuart Mill. Smith (1994) and Finlay (2014) defend more sophisticated versions. These views can provide a plausible metasemantic account for normative predicates if they can explain how the meaning-constituting descriptions pick out the relevant natural properties. The non-naturalists could attempt to avoid the previous problem by appealing to indirect reference-fixing descriptions like ‘... is the *sui generis* non-natural property that fills such and such role’ or ‘... is the *sui generis* non-natural property that is shared by most of these instances’. Here the question would be how the vocabulary in these definite descriptions manages to pick out their intended referents without relying on further descriptions. It could be suggested these words (such as ‘non-natural property’) acquire their semantic values through some form of ‘direct understanding’. However, if direct understanding is not sufficient to explain how normative predicates refer to the non-natural properties in the first place, then it is difficult to see how it could explain how the previous descriptions acquire their referents either.

new meaning-constituting descriptions contain only non-normative predicates in which case naturalism follows, or they include normative predicates in which case we face a regress again.

Non-naturalists are sometimes accused of ‘metasemantic quietism’, of ignoring the metasemantic questions (Wedgwood 2007: 3–4 and 21). Perhaps the non-naturalists have done so because they have assumed a more general form of descriptivism.<sup>13</sup> According to it, normative predicates refer to the relevant non-natural normative properties because (i) normative predicates refer to whatever properties they must in order for our carefully considered core normative convictions to be true and (ii) if these predicates referred to any other properties then at least some of those convictions would be false.

The non-naturalists could motivate (i) with the principle of charity.<sup>14</sup> It is widely accepted that, in trying to interpret what others mean, we should assume that they talk about objects, properties, and facts that make most of their utterances true. Many of the non-naturalists’ arguments against different naturalist reductions of normative properties could then be understood as supporting (ii). They attempt to show that, if the normative predicates referred to the natural properties to which the naturalists claim they refer, some of our core normative convictions would be false.<sup>15</sup>

This proposal fails for two reasons. Firstly, it fails to address the Contrast Challenge. §1 explained how, for every non-natural normative property posited by the non-naturalists, there is a necessarily co-instantiated natural

---

<sup>13</sup> See Lewis (1984: 224) and Sider (2011: 24). The first type of descriptivism above is ‘local descriptivism’. It assumes that the meaning-constituting descriptions already refer non-problematically. If, however, the meaning of all expressions consisted of some descriptions, we would face the problem of how our language manages to refer as a whole. ‘Global descriptivism’ is designed to address that question. This view is compatible with the idea that normative predicates are unanalyzable.

<sup>14</sup> See Dennett (1982), Lewis (1974), Quine (1960), and Davidson (2001).

<sup>15</sup> See, for example, Parfit’s (2011, vol. 1: 73–82) agony argument according to which analytic reductions of reasons cannot guarantee that we have reasons to avoid future agony.



property. The problem is that exactly the same first-order normative claims would be true even when the normative predicates referred to those natural properties. After all, whichever necessarily co-instantiated non-natural and natural property pair we take, exactly the same objects have those properties.

Secondly, it is worthwhile to recall that Hilary Putnam used his model-theoretic argument to show that we cannot use the truth of our first-order theories to explain reference in the previous way because doing so leaves reference 'inscrutable' (Putnam 1977 and 1981: 32–35 and 217–218).<sup>16</sup> In this context, the crux of his argument is that, whatever first-order normative theory we accept, every claim of that theory remains true when we systematically re-assign new referents to all the terms used to formulate our theory in a structurally isomorphic pattern. The sentence 'eating meat is wrong' thus remains true in the interpretation in which 'eating meat' refers to trench coats and 'is wrong' to being fashionable. We would, of course, then need to give a new referent to 'killing' too – such as holding a dinner party – given that the sentence 'killing is wrong' is true. If we continue re-assigning referents in this way, all the normative claims we currently take to be true will remain true. This is why the truth of our first-order normative theory is not sufficient to restrict the reference of the normative predicates to the intended non-natural properties.

The non-naturalists cannot adopt Fregean accounts of predicates either.<sup>17</sup> According to them, there are two kinds

---

<sup>16</sup> See also, for example, Williams (2007: sec. 1.2), Sider (2011: 24–27), and Button (2013: sec. 2.3). It is true that this argument applies more generally also in the non-normative contexts and so we might think that Putnam's argument does not pose a special problem for the non-naturalists who are global descriptivists. However, for the sake of what will follow below, it is useful to introduce the problem here. The naturalists might think that they can use the causal powers of natural properties to overcome the model-theoretic objection. In §4, I will argue that this attempt will fail and then in §5 I will argue that the non-naturalists can give the same response to the objection as the naturalists.

<sup>17</sup> See Frege 1997b [1891]. Even if this view is discussed in this section, it is not a form of descriptivism. Frege's idea was that the logical qualities of the predicates (mainly their unsaturatedness) are sufficient

of expressions. Names are ‘complete’ expressions because they refer to objects. In addition to them, there are also ‘unsaturated’ expressions: predicates with empty places that can be filled up by names. Frege then understood the latter expressions with the analogy of mathematical functions. The idea is that the predicates refer to functions from objects to truth-values (True and False), which Frege thought to be the referents of the standard subject-predicate sentences (Frege 1997b [1891]: 139–140).<sup>18</sup>

This view, unfortunately, does not leave room for metaethical non-naturalism. Because for every assumed non-natural normative property there is a necessarily co-instantiated natural property (see §1), there is also a necessarily co-extensive non-normative predicate for every normative predicate. Such predicates are long disjunctive predicates in which each disjunct consists of a long conjunction formulated in terms of the non-normative predicates that wholly describe the natural properties of a given object. So, for ‘...is good’ there is a necessarily co-extensive predicate of the form ‘... is D<sub>1</sub>, D<sub>2</sub>, ..., or D<sub>n</sub>’.

The problem is that functions are extensional.<sup>19</sup> If two seemingly different functions have the same extension for every argument, they are the same function. If we in this situation understand predicates in the Fregean way as functions from arguments to truth-values, then the two predicates in the necessarily co-extensive normative and non-normative predicate pairs will refer always to the same function. As a consequence, the Fregean understanding of predicates leaves no room for the non-naturalists to argue that normative predicates refer to something different (the *sui generis* normative properties) than to what the naturalists claim they refer.

---

to fix their reference to functions. I argue below that this metasemantic account is not available for the non-naturalists because of the kind of semantic values it assigns for predicates.

<sup>18</sup> Frege called functions whose value is always a truth-value ‘concepts’, which is why he thought that predicates refer to concepts, i.e., functions whose value is either True or False.

<sup>19</sup> See Miller (2007: 16). Frege himself, however, might have resisted this claim (Dummett 1981: 209).

### 3. Conceptual Role Semantics

Ralph Wedgwood (2001 and 2007, ch. 4) has suggested that the non-naturalists too could use conceptual role semantics to explain how normative predicates acquire their reference.<sup>20</sup> Its starting point is that there are basic rules of use that determine both when each word of our language can be used rationally and what a speaker must understand in order to be competent with the meaning of a given word. Wedgwood then argues that these rules are sufficient to determine the reference of each word.<sup>21</sup> The reference of a given word is roughly the object or a property that best makes sense of the basic rule that governs its rational use (Wedgwood 2001: 10).

Take the logical connective ‘or’. According to Wedgwood, being competent with its meaning requires mastering two basic rules: (i) accepting <P> commits you to accepting <P or Q> and (ii) accepting <Q> commits you to accepting <P or Q> (Wedgwood 2001: 7). Their mastery consists of having a disposition to follow them in reasoning and of finding the transitions that comply with them obviously correct and in no need of further justification.

Wedgwood then offers two principles of interpretation, *validity* and *completeness*, which take us from the basic rules to the referents (2001: 10–12 and 2007: 86–87). We must assign ‘or’ such a referent that in virtue of it (i) each instance of applying the previous two rules is ‘valid’ and (ii) the basic rule provides a ‘complete’ specification of the referent (Wedgwood 2001: 10). (i) entails that the referent must make it necessary that, if the input state to an inference according to the previous rules (belief that P, for

---

<sup>20</sup> As Wedgwood (2001, 7, fn. 12) acknowledges, his conceptual role semantics owes much to Christopher Peacocke’s (1987 and 1992: 1–40) account of logical constants. David Enoch (2011: sec. 7.6) has adopted Wedgwood’s semantic account (with minor amendments) to provide a non-naturalist account of reference. Wedgwood’s (2001 and 2007: ch. 7) own formulation of the view is technical. Schroeter and Schroeter (2003: secs. 2–4) and Merli (2009: sec. 2) are helpful.

<sup>21</sup> Here Wedgwood follows Block (1987) and Peacocke (1987 and 1992: ch. 1). Harman (1999), in contrast, believes that conceptual role semantics conflicts with truth-conditional semantics, whereas Field (1977) takes the two elements of meaning, conceptual role and reference, to be independent of one another.

example) is true, then the conclusion state (belief that P or Q) is true too. (ii), in contrast, entails that the referent of 'or' must make it impossible that the belief that P or Q is true except when the belief that P is true or the belief that Q is true.<sup>22</sup> Wedgwood then claims that, because of these constraints set by the previous two rules of interpretation, the referent of 'or' must be the truth function of the classical disjunction as it is defined by the standard truth tables (Wedgwood 2001: 11).

We must make a minor adjustment to the previous theory, because the relevant meaning-constituting basic rules of competence for normative predicates can contain practical mental states such as intentions, preferences, and plans. For example, the basic rule of competency for 'better than' is according to Wedgwood roughly the following: the rational acceptance of <X is better than Y> commits you to prefer X to Y (Wedgwood 2001: 15). Here the interpretation rules cannot be formulated in terms of truth-preservation because the output state, a preference, lacks a truth-value.

Wedgwood therefore reformulates the interpretation rules in terms of more general correctness (Wedgwood 2001: 18): the referents must (i) make following the basic rules 'correctness preserving' (validity) and (ii) the output states of those rules correct only if the input states are correct (completeness). He then needs to explain what it is for preferences and other practical states to be correct in the same way as beliefs are correct when they are true. Here Wedgwood believes that it is enough to think that these states are correct when they conform to the goals of practical reasoning, whatever they may be (Wedgwood 2001: 19).

---

<sup>22</sup> The basic rule captures the conditions in which it is a mistake to reject the conclusion state. The crux of completeness is that the meaning-constituting basic rules must capture all such conditions (Wedgwood 2001: 10–11). A set of rules does so when the conditions built into it are the only conditions in which rejecting the conclusion would be a mistake. Thus, the weaker referent 'p or q or r' would make inferences in accordance with the introduction rule for disjunction correct (and so this referent would satisfy validity) but it would not satisfy completeness because rejecting p or q is not a mistake when one believes that only r is true.

To what property does this account then fix the reference of 'better than'? Wedgwood claims that this predicate must refer to the normative property of betterness, that is, the property of *being what it is practically correct to prefer given the goals of practical reasoning, whatever they are*, because only then it would be the case both that:

- necessarily, if the input to using the previous basic rule is correct (the belief that X is better than Y is true), then preferring X to Y is correct (i.e., it conforms to the goals of practical reasoning, whatever they are) [and]
- necessarily, preferring X to Y is correct (i.e., it conforms to the goals of practical reasoning, whatever they are) only if the belief that that X is better than Y is true.

The idea is that, if 'better than' referred to any other property, then the previous conditionals would fail in some cases. That, however, just cannot happen because they capture the basic general rules of interpretation as they apply to the predicate in question.

The problem of this metasemantic account in the non-naturalists framework is that it cannot meet the Contrast Challenge.<sup>23</sup> This is because the basic rule of competence for 'better than' and the two interpretation rules of validity and completeness are too sparse to fix the reference of this predicate uniquely to a non-natural property.

According to Wedgwood, betterness is the property of being what it is correct to prefer given the goals of practical reasoning, whatever they are. This is because he believes that only if the latter property is the referent of 'better than' validity and completeness will be respected. This means that the non-naturalists can adopt Wedgwood's theory only if they think that the normative property of being what it is practically correct to prefer given the goals of practical reasoning, whatever they are,

---

<sup>23</sup> For alternative criticisms, see Schroeter and Schroeter (2003) and Merli (2009). Schroeter and Schroeter argue that Wedgwood's conceptual roles for normative predicates either fail to rule out obviously non-normative referents or they presuppose a prior understanding of normative notions. Merli, in contrast, argues that Wedgwood's account cannot be applied to moral normative notions.

is a *sui generis non-natural property*. The problem is that Wedgwood's conceptual role account fails to fix the reference of 'is better than' uniquely to that kind of a property.

In §1 we saw that there is a disjunctive natural property that would be necessarily co-instantiated with the assumed non-natural property of being what it is correct to prefer given the goals of practical reasoning, whatever they are. The worry is that, if 'better than' referred to this natural property, this too would (i) make all the instances of inferring in accordance to the meaning-constituting basic rule of competence for 'better than' valid and (ii) it would also enable that basic rule to provide a complete characterisation of the referent.

According to this alternative interpretation, the input state to the rule – the belief that X is better than Y – attributes to preferring X to Y the natural property of being  $D_1, D_2, \dots$ , or  $D_n$ . When this belief is true, preferring X to Y will have exactly that natural property. This entails that inferring according to the rule 'Acceptance of <X is better than Y> commits you to prefer X to Y' will always be valid in Wedgwood's sense. Necessarily, if the belief that constitutes the input to following this rule is true under the current assignment of reference (i.e., preferring X to Y has the natural property of being  $D_1, D_2, \dots$ , or  $D_n$ ), then it will be correct to prefer X to Y given the goals of practical reasoning, whatever they are. This is because doing so will have the property of being  $D_1, D_2, \dots$ , or  $D_n$  and it and being correct given the goals of practical reasoning, whatever they are, are necessarily co-instantiated properties.

The same goes for completeness. Completeness requires that, necessarily, preferring X to Y is correct (that is, it conforms to the goals of practical reasoning, whatever they are) only if the belief that X is better than Y is true. Or, in other words, it is impossible for the given preference to be correct given the goals of practical reasoning, whatever they are, except when the latter belief is true. Let us assume both (i) that the property of being  $D_1, D_2, \dots$ , or  $D_n$  is necessarily co-instantiated with the property of conforming to the goals of practical reasoning,

whatever they are, and (ii) that it is the referent of the predicate ‘...is better than...’. In this situation, it cannot be that preferring X to Y is correct and yet the belief according to which it has the property of being  $D_1, D_2, \dots$ , or  $D_n$  is false. After all, given necessary co-instantiation, the correctness of a preference entails that it has the property of being  $D_1, D_2, \dots$ , or  $D_n$ .

This means that the constraints provided by the two interpretation principles of validity and completeness leave it open whether ‘better than’ refers to the non-natural property posited by the non-naturalists or to the natural property that is necessarily co-instantiated with it. Both potential referents satisfy the interpretation principles equally well. Because of this, Wedgwood’s conceptual role semantics fails to provide the non-naturalists with a solution to the Contrast Challenge.<sup>24</sup> We have been given no explanation of why the normative predicates would refer to the non-natural properties rather than to the underlying natural properties.

#### 4. Synthetic Naturalism and the Qua-Problem

It could then be thought that different forms of synthetic naturalism will have a theoretical advantage: unlike the non-naturalists, their defenders can use the essential qualities of natural properties to explain reference (Dunaway and McPherson 2016: 1). Synthetic naturalists believe that normative properties are ordinary natural properties and thereby causally efficacious. Perhaps they could therefore rely on causal connections between the core uses of normative predicates and natural properties to explain how the normative predicates manage to refer.

---

<sup>24</sup> Wedgwood might endorse this result. He notes that the interpretation rules leave it open whether ‘or’ refers to the familiar truth function of disjunction or to the function that maps any two sets of possible worlds onto their union (Wedgwood 2001: 11). This is because both functions, given their isomorphic extensions, make sense of the basic rule of use for ‘or’ equally well. He also admits that, if the goal of practical reasoning can be described entirely in naturalist vocabulary, then the conceptual role semantic account of ‘... is better than...’ will be compatible with reductive naturalism (Wedgwood (2001: 20).

Let us begin from a well-known naturalist causal theory of how normative predicates acquire their reference (Boyd 1988: especially 195).<sup>25</sup> On this view, the uses of ‘good’ in our linguistic community are causally regulated by certain natural properties in a certain epistemically privileged way. When we causally interact with these properties, our assertions about what is good come to be more frequently true over time – under the assumption that we are referring to those properties. The claim then is that the reference of ‘good’ is directly fixed to the natural properties that causally regulate the core uses of this term in the previous reliability-conducive way. Boyd (1988: sec. 4.3) also argues that, in the framework of this theory, we should think that ‘good’, for example, refers to a certain consequentialist homeostatic property cluster related to central human needs.

For the present purposes, it does not matter which causal mechanism is argued to fix the reference of normative predicates or to which natural properties that mechanism is claimed to fix their reference. This is because all causal theories of reference face the same decisive objection, the so-called *qua-problem*.<sup>26</sup>

The natural properties that causally regulate the core uses of normative predicates can only do so by causing token

---

<sup>25</sup> See also Schroeter & Schroeter (2013: sec. 3) and van Roojen (2015: sec. 11.2). Boyd’s account was inspired by the direct reference theories of Kripke (1980) and Putnam (1975). Other synthetic naturalists have developed improved versions of Boyd’s account (see, e.g., Brink (2001: 167–170)). Also other types of causal theories of reference are available for the synthetic naturalists. Janice Dowell (2016), for example, applies Ruth Millikan’s (1984) teleological metasemantic account to normative predicates. On this view roughly, what property a predicate refers to is determined by which mapping has contributed in the normal circumstances causally to the selection of the ‘consumer systems’ of the audience. See fn. 28 below.

<sup>26</sup> See, for example, Papineau (1979, sec. 5.7) and Devitt & Sterelny (1987, secs. 4.4 and 5.3). For a modern version, see Sider (2011: 34–35). The standard objection to causal theories of reference in metaethics is the Moral Twin-Earth objection (Horgan and Timmons 1991, 1992a and 1992b). For responses, see Copp (2000); Dowell (2016); Dunaway and McPherson (2016); Laurence, Margolis, and Dawson (1999); Merli (2002); and van Roojen (2006).



uses. The synthetic naturalists must thus have in mind the following picture:

Context <sub>1</sub>	Context <sub>2</sub>	Context <sub>3</sub>	Context <sub>4</sub>	...
Use <sub>1</sub>	Use <sub>2</sub>	Use <sub>3</sub>	Use <sub>4</sub>	...
↑	↑	↑	↑	
NP <sub>1</sub> I <sub>1</sub>	NP <sub>1</sub> I <sub>2</sub>	NP <sub>1</sub> I <sub>3</sub>	NP <sub>1</sub> I <sub>4</sub>	...

Here NP<sub>x</sub>I<sub>y</sub> stands for a token y of the instantiations of the natural property (NP<sub>x</sub>), which is assumed to causally regulate our uses of 'good' in the right way and which thus is its alleged referent. This property does the causal regulation by causing the core uses of 'good' in every normal context of use. The arrows then represent the causal relations between the relevant instantiations of the natural property and the token uses of the predicate.

The instantiations of the relevant natural property, however, never directly cause 'good' to be used but rather this always happens through causal chains. First the instantiations of the natural property, which the naturalists take to regulate our uses of 'good', are caused by the instantiations of some other natural properties. The instantiations of that property then cause various other things to happen – such as certain processes in our brains – which finally cause us to use the predicate 'good' in thought and speech. The picture should thus be more like this:

Context <sub>1</sub>	Context <sub>2</sub>	Context <sub>3</sub>	Context <sub>4</sub>	...
Use <sub>1</sub>	Use <sub>2</sub>	Use <sub>3</sub>	Use <sub>4</sub>	...
↑	↑	↑	↑	
NP <sub>4</sub> I <sub>1</sub>	NP <sub>7</sub> I <sub>1</sub>	NP <sub>14</sub> I <sub>1</sub>	NP <sub>2</sub> I <sub>1</sub>	...
↑	↑	↑	↑	
NP <sub>1</sub> I <sub>1</sub>	NP <sub>1</sub> I <sub>2</sub>	NP <sub>1</sub> I <sub>3</sub>	NP <sub>1</sub> I <sub>4</sub>	...
↑	↑	↑	↑	
NP <sub>23</sub> I <sub>1</sub>	NP <sub>14</sub> I <sub>2</sub>	NP <sub>5</sub> I <sub>1</sub>	NP <sub>104</sub> I <sub>1</sub>	...

This table represents the idea that, in the relevant causal chains, there is always one instantiation of a different natural property between the instantiations of the property that allegedly is the referent of 'good' and the relevant uses of the predicate. It also suggests that there is one instantiation of another natural property in these causal chains before the instantiation of the natural property that is the alleged referent of 'good'. In these causal chains there will, however, always be many instantiations of many different natural properties both between the core uses and the instantiations of the alleged referent natural property and also before the relevant instantiations of that natural property.

As a consequence, there will be many disjunctive properties that are always present in the relevant causal chains that connect the instantiations of the natural property that is claimed to be the referent of 'good' to the core uses of 'good'. In the previous table, one such property would be the property of being NP<sub>4</sub>-in-Context<sub>1</sub>-or-NP<sub>7</sub>-in-Context<sub>2</sub>-or-NP<sub>14</sub>-in-Context<sub>3</sub>-or-NP<sub>2</sub>-in-Context<sub>4</sub>-or-...-or-NP<sub>11</sub>-in-Context<sub>n</sub>. By selecting instantiations of different natural properties from the relevant causal chains in this way, we can pick out uncountably many new gerrymandered disjunctive properties that are part of all the same causal chains with the core uses of 'good' as NP<sub>1</sub>.<sup>27</sup>

The qua-problem is the concern that, if NP<sub>1</sub> is causally connected to the core uses of 'good' in the way specified by the naturalists' causal metasemantic theory, then all the previous disjunctive properties will be causally connected to the same core uses exactly in the same way. Hence, if NP<sub>1</sub> is claimed to causally regulate the core uses in the right way, the constructed properties must do so too. This

---

<sup>27</sup> We might be even able to form a single aggregated property from all the disjuncts in all the relevant causal chains, which would leave us with a single horribly gerrymandered natural property. The problem is that the causal theories of reference as such leave it open whether the normative predicates would refer to this type of properties, the original natural properties that the naturalists take to be the referents, or the other gerrymandered natural properties we can construct that are always co-present.

leaves the reference of ‘good’ underdetermined. Given the constraints set by the synthetic naturalists’ causal theories of reference, the referent of ‘good’ could be any one of the natural properties that are a part of the very same causal chains. For this reason, no amount of theorizing about causal connections will solve the problem.<sup>28</sup>

## 5. Reference Magnetism and Non-Naturalism

In the previous argument, nothing turns on the fact that we are discussing normative predicates. The qua-problem is, after all, a general challenge for all causal-historic theories of reference no matter to what expressions they are applied. The synthetic naturalists should therefore begin from how the problem can be solved elsewhere.

The most promising solution to the qua-problem relies on ‘reference magnetism’.<sup>29</sup> On this view, properties are on a spectrum with respect to how similar having them makes their bearers. It is a brute fact that some properties are *elite*: having them constitutes fundamental structural similarities between objects. Metaphorically speaking,

---

<sup>28</sup> This shows how causal relations are less fine-grained than semantic ones (Loewer 1997: 112). It could be responded that more sophisticated causal theories of reference will have resources to deal with this problem. On the behalf of Boyd it could be suggested that, even if the gerrymandered properties are part of the same causal chains as the intended referents, they do not causally *explain* the relevant uses of ‘good’ as well. Likewise, drawing from Millikan, it could be argued that the gerrymandered properties fail to explain for which mapping relation the relevant consumer systems are selected in evolution. As Sider (2011: 30) notes, the relevant question then is: why would the gerrymandered properties provide worse evolutionary explanations or worse explanations of our core uses? The intended referents seem to offer better explanations precisely because they are less gerrymandered – they are more brute similarities in nature. This is why the natural defenses of Boyd and Millikan lead to the metasemantic framework discussed in the next section.

<sup>29</sup> This theory was first proposed by G.H. Merrill (1980). It is standardly attributed to David Lewis (1983, 1984, and 1986), even if Schwarz (2014) argues that it wasn’t Lewis’s view. For recent defenses, see Brian Weatherson (2003) and Ted Sider (2011: sec. 3.2). It is sometimes suggested that the referential intentions of the speakers suffice to solve the qua-problem. Such solutions, however, fail to explain how the referential intentions manage to refer to the salient properties (Lewis 1984, 226).

these properties constitute nature's joints. For example, the property of being gold is an elite property, because its bearers must have the atomic number 79. This microphysical property constitutes a fundamental similarity between the objects that have the property in question. In contrast, fashionableness is not elite: trench coats and dinner parties, for example, need not have much in common *qua* being fashionable. Properties are not, of course, either wholly elite or not elite at all but rather they can be more or less elite depending on how structurally similar objects are in virtue of having them.

Reference magnetism then is the view that the more elite a property is the more intrinsically eligible it is for being the referent for a given predicate: in other words, elite properties function as 'reference magnets'.<sup>30</sup> Hence, when the other constraints on the assignment of reference (such fitting linguistic usage, preserving the truth of our first-order theories, and making sense of inferential connections) leave many properties as eligible candidates, the genuine referent is the most elite of those properties. According to the reference magnetism, the reference of our predicates attaches more easily to such properties.<sup>31</sup>

---

<sup>30</sup> In addition, elite properties are also thought to feature in genuine laws and claims including them are claimed to be easier to confirm (Dunaway 2016, 249; Sider 2012: ch. 3).

<sup>31</sup> This solution may seem like an *ad hoc* response to the previous problem (Putnam 1981: 53). Reference magnetism can, however, be motivated by the fact that it offers a response to a range of semantic paradoxes such as Quine's (1960) argument for indeterminacy of translation, Putnam's permutation argument (§2), and the rule-following paradox (Kripke 1982) based on an attractive metaphysical theory (Dunaway and McPherson 2016: sec. 1.3). Secondly, semantic accounts can be used to explain linguistic behaviour. In this situation, reference magnetism can be motivated with the theoretical virtue of simplicity as it guides us to prefer explanations based on sparse properties (Williams 2007: sec. 2). Despite these advantages, the idea of reference magnetism remains controversial (see, e.g., Schwarz (2014) and Sundell (2012)). My conclusion is thus modest: if reference magnetism is a viable metasemantic view, then the non-naturalists can use it to explain how normative predicates refer to non-natural properties. This corresponds to Dowell's (2016) argumentative strategy; she argues that, if Millikan's teleosemantics – a controversial metasemantic position – is along the right lines, then the Moral Twin Earth objection to synthetic naturalism fails. There may well also be

Synthetic naturalists can use the previous theory to solve the qua-problem.<sup>32</sup> Boyd, for example, could argue that the consequentialist homeostatic property cluster related to human needs is the most elite candidate for being the referent of 'good' of all the natural properties that causally regulate the core uses of this predicate in the right way. After all, all the other properties that belong to the relevant causal chains are disjunctive properties constructed from the instantiations of different natural properties in the situations in which the consequentialist homeostatic property cluster is also present. Boyd could then claim that, because the latter property is the most elite candidate, it attracts the reference of 'good'.

There is, however, nothing causal about reference magnetism. The lesson of the qua-problem is that causation always leaves room for many eligible candidates when we consider to what a given word refers. If the more elite candidates are then to attract reference, they must do so in some *non-causal* way. Because of this, there is no antecedent reason to think that only natural properties – *given how naturalness is understood in metaethics in terms of causally efficacy, a posteriori knowability, and being a part of the best explanations of natural and social sciences* – can function as reference magnets.<sup>33</sup> This means that non-naturalists too should at least in principle be in a position to solve the metasemantic challenges facing them by relying on reference magnetism on the condition that they can argue that the non-natural properties are elite and thus function as reference magnets (see §5.1 below).

---

other metasemantic accounts that could help the non-naturalists to avoid the problem too (see, for example, Grice (1957) and Horwich (1998)). Here I only want to suggest that there is at least one such account.

<sup>32</sup> Synthetic naturalists have already relied on reference magnetism to respond to the Moral Twin-Earth objection (van Roojen 2006; Dunaway and McPherson 2016). Dunaway (2015 and 2016) has argued that the non-naturalists should rely on reference magnetism both to distinguish their view from forms of quasi-realism and to respond to Jackson's (1998) objection to non-naturalism.

<sup>33</sup> Lewis (1983: 346) defined naturalness in terms of eliteness, which makes the claim that only natural properties are reference magnets trivial. Metaethical non-naturalists can agree that normative properties are natural if naturalness is understood in this way.

Of course, a certain amount of modesty is called for at this point.<sup>34</sup> As explained in fn. 31, it is worth highlighting that reference magnetism itself remains a controversial metasemantic view. The aim of this section is therefore merely to make two suggestions. Firstly and as I have already argued, the metaethical naturalists too have at least some good reasons to rely on reference magnetism and thus they are likely to be ‘companions of guilt’. Secondly and as I will argue below, if the phenomenon of reference magnetism exists as many have thought and the non-natural properties are elite, then the non-naturalists seem to be in a position to rely on reference magnetism to solve the metasemantic challenges facing them.

To see how the non-naturalists might carry out this project, let us consider global descriptivism described in §2. According to it, normative predicates refer to whatever properties they must in order for our carefully considered normative convictions to be true. Let us assume that the non-naturalists then combine this metasemantic theory with reference magnetism – the view according to which, when there are many eligible candidates, the most elite attract reference.<sup>35</sup> We can then consider whether non-naturalists could use the previous combination of views to meet the metasemantic challenges introduced in §1. I will begin from the easiest challenge and then move onto the more difficult ones.

### 5.1 The Gruesome Normative Property Challenge

The non-naturalists have always believed that normative properties are simple, *sui generis*, non-reducible brute

---

<sup>34</sup> According to Dunaway, non-naturalists should rely on reference magnetism also to explain both (i) why there is no widespread indeterminacy in what normative predicates refer to and (ii) how communities who use these terms differently can still disagree (Dunaway 2016: sec. 10.3.1).

<sup>35</sup> This view follows Sider (2011: 31–33; see also Dunaway 2016: 252). For reasons explained by Sider (2011: 31), we should not understand reference magnetism to be a metasemantic view of its own (there are after all words that do not refer to elite properties) but rather as a ‘meta-metasemantic’ view, which may be combined with other metasemantic views. The metasemantic views tell us about the nature of reference whereas reference magnetism determines to which eligible candidate the reference-relations attach.

similarities that belong to the ‘fundamental nature of the universe’ (Ross 2002 [1930]: 29–30). One benefit of this metaphysical foundation of non-naturalism is that it makes the normative properties perfectly elite on the criterion of eliteness specified above.<sup>36</sup> It is natural to think that especially the *simplicity* of the posited non-natural normative properties makes these properties excellent candidates for being perfectly elite. When we take a less than perfectly elite property, the constitution of such properties can be understood to be analysable in terms of other, more basic properties exactly as Lewis (1984: 228) thought. When it comes to simple properties that have no other properties as their constituents, there just is no way of giving an account of these properties in terms of anything more basic. The having of such properties is a matter of being similar in a fundamental, brute way, which is exactly why it is appealing to think that the properties in question are perfectly elite.

Gruesome normative properties are, in contrast, less elite given that objects that have them need not have much in common (consider, for example, any good action inside our light cone and a cruel one outside it). Because more elite properties then attract reference more easily, the normative predicates, in virtue of their simplicity, can be argued to refer to the standard non-natural normative properties instead of their gruesome counterparts (see Lewis 1984: 228 and Sider 2011: 33).

---

<sup>36</sup> There are places where Lewis (1984: 228) defines the most fundamental microphysical properties as perfectly elite and then claims that how elite other properties are is a matter of the length of their canonical definitions in terms of the perfectly elite properties (Lewis 1986: 61). For why this view of eliteness fails, see Dunaway (2016: 251–252), Hawthorne (2006: 206 and 2007: 434), Schroeter & Schroeter (2013: 17), and Williams (2007). For this reason, I follow Dunaway and McPherson (2016: sec. 3.2) and Schroeter & Schroeter (2013: 18) in understanding degrees of eliteness as a primitive notion that refers to brute similarities in the world. For a discussion, see Dunaway (2016: 251). Dunaway (*ibid.*) also rightly points out that non-naturalists have not explicitly stated their views in terms of eliteness.

## 5.2 The Contrast Challenge

Can the non-naturalists use reference magnetism also to explain why normative predicates refer to non-natural properties instead of natural properties? Global descriptivism may at first seem to rule out most natural properties as potential referents of normative predicates. According to it, the referents of the normative predicates must make our carefully considered normative convictions true. Consider then all the objects that are good according to those convictions. Due to the previous interpretation rule, 'good' could only refer to a natural property that is had by all of them. However, given how many different kinds of actions, character-traits, and states of affairs we take to be good, the global descriptivist interpretation principle does not seem to leave many natural properties to be eligible candidates for being the referent of 'good' (McDowell 1998 [1981], 201–202).

We know, however, that at least one eligible natural candidate must remain. According to the non-naturalists, all objects that are good according to our carefully considered convictions have a certain non-natural property. If that property were the referent of 'good', those convictions would clearly be true. They would be ascribing a property to a set of objects, which all of them would have. Given the reasoning explained in §1, we know that there is a natural property that would be necessarily co-instantiated with the previous non-natural property. If that natural property were the referent of 'good', this too would make our carefully considered convictions equally true (§2).

Here the non-naturalists can rely on reference magnetism to explain why 'good' would refer to the non-natural property rather than to the necessarily co-instantiated natural property. The thought is that 'good' refers to the non-natural property because it is more elite than the natural property in question (Dunaway 2015: 650). As already explained in §5.1, the non-naturalists' metaphysical theory makes the non-natural normative properties perfectly elite, whereas the necessarily co-instantiated natural property could hardly be as elite. That property is, after all, a disjunctive property in which each



disjunct consists of all the natural properties of a given object that is good according to our carefully considered convictions.

The problem is that different objects that have the underlying disjunctive natural property are not very similar to one another *qua* having that property. Take, for example, a just global distribution of resources and Mary's act of helping her elderly aunt with her shopping, which are both good things. These objects are, however, not very similar to one another in virtue of *all the natural properties they have*. The former state of affairs affects billions across the globe, whereas the latter concerns only two individuals locally. Because of this, the non-naturalists can argue that, of the eligible candidates, the relevant non-natural properties are more elite than the underlying necessarily co-instantiated natural properties, and therefore they must be the referents of the normative predicates.<sup>37</sup>

Lewis (1984) argued that reference magnetism also blocks Putnam's permutation-challenge. §2 explained how our best first-order normative theory remains true even when we re-assign new referents to all the expressions used to formulate that theory. So, 'eating meat is wrong' remains true when 'eating meat' refers to trench coats and 'is

---

<sup>37</sup> It could be objected that the disjunctive natural property need not be the only natural property shared by all good things (perhaps the disjunctiveness is merely a matter of the logical complexity of the predicate in question rather than a genuine feature of the world). It could be, for example, thought that the shared natural property must be an empirical dimension of similarity on which the normative properties supervene – a pattern relevant from the point of view of normative thinking (Jackson, Smith, and Pettit 2000). Good things could thus also have, for example, the property of maximizing the amount of pleasure or the property of being the object of consistent rational willing, and they certainly have the sociological property of being conventionally good. Given that all the previous natural properties are presumably less elite than the *perfectly elite* natural properties such as having a negative electric charge, it is not obvious that they would be as elite as the equally perfectly elite simple non-natural properties posited by the non-naturalists. It is true, however, that the non-naturalists will need to be able to explain at this point how, for the purposes of the relevant comparisons, we are to evaluate how elite the allegedly less than perfectly natural properties are (see Schroeter & Schroeter (2013: 18–20)).

wrong' to being fashionable. If we keep reassigning referents like this in a systematic, structurally isomorphic way (§2), then no part of our first-order theory becomes false even with the crazy new referents. Given the constraints set by global descriptivism and the possibility of this type of permutations, the normative predicates could thus refer to any natural properties.

Lewis's insight was that, even if global descriptivism fails to rule out the crazy new assignments of reference, reference magnetism can do so (Lewis 1984: 227). Consider the previous example. According to the non-naturalists' original interpretation, a certain simple non-natural property is the referent of 'good'. Given the possibility of permutations, global descriptivism, however, means that this predicate could equally well refer to any natural property as we can find suitable new referents to all other words to make our first-order normative theory true even in that case. Reference magnetism, however, rules out all the new re-assignments that give 'good' a natural referent that is less elite than the initial non-natural property.

According to the non-naturalists, the non-natural property of goodness is perfectly elite: it stands for a fundamental brute similarity in the universe – it carves the universe at its joints. This means that the non-naturalists can reasonably rule out all the less than perfectly elite candidate referents that are acquired through the permutations. They can exclude not only fashionableness but also almost all other natural properties too (fn. 37). In fact, perhaps the only eligible equally perfectly elite natural property candidates would be the fundamental physical properties of mass, charge, quark colour and flavour, and the like (Lewis 1984: 228, but see fn. 35 above).

Furthermore, even the idea that normative predicates could refer to the most fundamental microphysical properties is problematic. In order to make our best first-order normative theory true, we would need to assign a different microphysical property as a referent for every normative predicate given that our best first-order normative theory gives different extensions for different normative properties. This means that a successful

reassignment of referents in a permutation would require that there were exactly as many fundamental microphysical properties as there are normative properties according to our best first-order normative theory.<sup>38</sup> Given how unlikely this is, normative predicates probably cannot be permuted to refer to the fundamental microphysical properties either.

This means that the non-naturalists seem able to exclude the permutations that assign either perfectly elite natural properties or other less elite natural properties to be the referents of normative predicates.<sup>39</sup> Because of this, the non-naturalists have reason for optimism: the combination of global descriptivism and reference magnetism seems to help them to respond to the Contrast Challenge.

### 5.3 The Correct Property Challenge

This leaves us with the most difficult challenge: why does 'good', for example, refer to the non-natural property goodness rather than to any other non-natural normative property? Here the non-naturalists cannot rely on reference magnetism because presumably all non-natural normative properties are equally elite. The non-naturalists might perhaps suggest that 'good' could not refer to

---

<sup>38</sup> That this possibility exists illustrates how there are ways the world could be in which reference magnetism would not solve Putnam's paradox. Williams (2007: sec. 3) offers an account of precisely when the conditions for this are met. He (2007: sec. 4) also explains why we might actually be in circumstances in which reference magnetism does not rule out that our ordinary expressions have crazy referents such as numbers. For a further development of Williams's concerns, see Hawthorne (2007). If reference magnetism fails to solve the general semantic paradoxes, then both naturalists and non-naturalists face exactly the same metasemantic problems.

<sup>39</sup> It could also be objected that, assuming global descriptivism and reference magnetism, some of our ordinary non-normative predicates could end up referring to the non-natural normative properties given that such properties would be perfectly elite and thus able to attract the reference of these terms if the use of these terms at least roughly fitted this interpretation (see Sundell (2012), Schroeter & Schroeter (2013), and Williams (2015)). However, if our ordinary non-normative predicates referred to the non-natural properties, then we would be unable to know whether they applied to an actual object *a posteriori*. If 'red' thus referred to a non-natural property, then we would be unable to know whether an object is red just by looking at it.

wrongness, for instance, because this would make many of our core normative convictions false. If 'good' referred to wrongness, a just distribution of resources could not, for example, be good.

This response, however, takes us right back to Putnam's paradox. The problem is that, if global descriptivism were the only constraint on interpretation, this would allow us to give new referents also for all the expressions that are used to refer to different actions, character traits, states of affairs, and the like. With a suitable re-assignment of referents for these expressions (which also does not give them less elite referents), our normative convictions would remain true even if our normative predicates referred to different non-natural normative properties than usually thought. The combination of global descriptivism and reference magnetism thus fails to solve the most difficult metasemantic challenge.

Let me finish by outlining how the non-naturalists could respond to this final objection by drawing from moral functionalism.<sup>40</sup> I have so far referred loosely to our best first-order normative theory, which I have assumed to consist of our carefully considered normative convictions about what one ought to do in different situations, what reasons there are for different actions in them, what is good and bad, and so on. Yet, our best first-order normative theory can also be understood to contain a set of platitudes that connect different normative predicates to one another. These platitudes make explicit defeasible inferential connections between different predicates, which their competent users must master (see §3 above and Jackson and Pettit (1995: 22–24)). Jackson, for example, gives the following examples this type of platitudes: 'courageous people are more likely to do what is right than cowardly people', 'the best option is the right option', and 'rights impose duties' (Jackson 1998: 130–131).

The non-naturalists could then agree with Jackson and Pettit (1995: 24–25) that one central element of the meaning of a given normative predicate is the place it has

---

<sup>40</sup> See Jackson and Pettit (1995) and Jackson (1998: ch. 5).

in the network constituted by the previous interconnecting platitudes. This element of meaning then provides an additional constraint, which the correct assignment of reference must satisfy. The idea is roughly that the place of a given normative predicate in the network provides a 'job description', which a property must fulfil if it is to be an eligible candidate for being the referent of the predicate (Schroeter and Schroeter 2003: 191). This means that the normative property to which, say, 'good' refers must be related to other normative properties in the way in which the network of the interconnecting platitudes relates the property to other properties in our best first-order normative theory.

The non-naturalists could then use this additional metasemantic constraint to explain why 'good', for example, must refer to goodness instead of any other normative property. For this to succeed, the interconnecting network of platitudes must have one important feature: it must be asymmetric. The network of relations ascribed to a given normative property by the interconnecting platitudes must be unique: structurally different from how all other properties in the network relate to each other.<sup>41</sup> If this condition is satisfied, then the account will be able to explain why 'good' can only refer to goodness rather than to any other non-natural normative property.

This is because, if 'good' in this situation referred to any other property, that property would not be related to the other normative properties in a structurally isomorphic way to how goodness relates to other properties. As a consequence, at least some of the interconnecting platitudes in which the predicate 'good' features would be false, no matter how we tried to reassign different normative properties to be the referents of the other normative predicates. This is why the network understanding of the meaning of normative predicates promises to help the non-naturalists to solve the Correct

---

<sup>41</sup> See Smith (1994: 48–54). Smith (1994: 54–56) is skeptical about whether there is enough substance in the platitudes to provide the necessary asymmetry. James Lenman (2006: sec. 4.3) is more optimistic.

Property Challenge – as long as our best first-order normative theory connects different normative predicates to each other in structurally asymmetric ways.

## 6. Conclusion

In conclusion then, the non-naturalists appear to be in a position to explain how the normative predicates manage to refer to the non-natural properties as they are characterized by the non-naturalists core metaphysical commitments. By adopting a combination of global descriptivism, reference magnetism, and a network account of the meaning of normative predicates, the non-naturalists are in a position to explain (i) why normative predicates do not refer to gruesome properties, (ii) why they refer to non-natural properties rather than to natural properties, and (iii) why the non-natural properties refer to the non-natural properties they do rather than to some other non-natural properties.

This is significant because, in addition to the already well-known metaphysical assumptions of the non-naturalists, the previous metasemantic explanations rely only on elements of metasemantic views, which are already widely accepted by many naturalists. This means that, other than objecting to their metaphysical commitments in the old-fashioned way, the naturalists cannot have any additional objections to non-naturalism on metasemantic grounds. Both metaethical naturalists and non-naturalists must rely on the very same theories in metasemantics, and so we must decide between the views on some other grounds.

Jussi Suikkanen  
University of Birmingham  
Department of Philosophy  
j.v.suikkanen@bham.ac.uk

## References

Bedke, Matthew (2009): 'Intuitive Non-Naturalism Meets Cosmic Coincidence'. *Pacific Philosophical Quarterly*, 90: 188–209.

- Block, Ned (1987): 'Functional Role and Truth Conditions'. *Proceedings of the Aristotelian Society*, suppl. 61: 157–181.
- Boyd, Richard (1988): 'How to Be a Moral Realist'. In G. Sayre-McCord (ed.): *Essays on Moral Realism* (Ithaca, NY: Cornell University Press), 181–228.
- Brink, David (1989): *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press).
- Brink, David (2001): 'Realism, Naturalism, and Moral Semantics'. *Social Philosophy & Policy*, 18: 154–176.
- Button, Timothy (2013): *The Limits of Realism* (Oxford: Oxford University Press).
- Copp, David (2000): 'Milk, Honey, and the Good Life on Moral Twin Earth'. *Synthese*, 124: 113–137.
- Copp, David (2003): 'Why Naturalism?' *Ethical Theory and Moral Practice*, 6: 179–200.
- Davidson, Donald (2001): *Inquiries into Truth and Interpretation* (Oxford: Oxford University Press).
- Dennett, Daniel (1982): 'Making Sense of Ourselves'. In J. I. Biro & R. W. Shahan (eds.): *Mind, Brain, and Function: Essays in the Philosophy of Mind* (Norman, OK: University of Oklahoma Press), 63–81.
- Devitt, Michael and Sterelny, Kim (1987): *Language and Reality – an Introduction to the Philosophy of Language* (Oxford: Basil Blackwell).
- Dowell, Janice (2016): 'The Metaethical Insignificance of Moral Twin Earth'. *Oxford Studies in Metaethics*, 11: 1–27.
- Dummett, Michael (1981): *Frege: Philosophy of Language*, 2<sup>nd</sup> edition (London: Duckworth).
- Dunaway, Billy (2015): 'Supervenience Arguments and Normative Non-Naturalism'. *Philosophy and Phenomenological Research*, 41: 627–655.

- Dunaway, Billy (2016): 'Expressivism and Normative Metaphysics'. *Oxford Studies in Metaethics*, 11: 241–265.
- Dunaway, Billy and McPherson, Tristram (2016): 'Reference Magnetism as a Solution to the Moral Twin Earth Problem'. Forthcoming in *Ergo*.
- Enoch, David (2011): *Taking Morality Seriously* (Oxford: Oxford University Press).
- Field, Hartry (1977): 'Logic, Meaning, and Conceptual Role'. *Journal of Philosophy*, 69: 379–408.
- Finlay, Stephen (2014): *Confusion of Tongues – A Theory of Normative Language* (Oxford: Oxford University Press).
- Frege, Gottlob (1997a [1892]): 'On *Sinn* and *Bedeutung*'. In M. Beaney (ed.), *The Frege Reader* (Oxford: Blackwell), 151–171.
- Frege, Gottlob (1997b [1891]): 'Function and Concept'. In M. Beaney (ed.), *The Frege Reader* (Oxford: Blackwell), 130–148.
- Gibbard, Allan (2003): *Thinking How to Live* (Cambridge, MA: Harvard University Press).
- Goodman, Nelson (1983): *Fact, Fiction, and Forecast*, 4<sup>th</sup> edition (Cambridge, MA: Harvard University Press).
- Grice, Paul (1957): 'Meaning'. *Philosophical Review*, 66: 377–378.
- Harman, Gilbert (1999): '(Non-Solipsistic) Conceptual Role Semantics'. In his *Reasoning, Meaning, and Mind* (Oxford: Oxford University Press), 206–232.
- Hawthorne, John (2006): 'Epistemicism and Semantic Plasticity'. In his *Metaphysical Essays* (Oxford: Oxford University Press), 185–210.
- Hawthorne, John (2007): 'Craziness and Metasemantics'. *The Philosophical Review*, 116: 427–441.



- Horgan, Terence and Timmons, Mark (1991): 'New Wave Moral Realism Meets Moral Twin-Earth'. *Journal of Philosophical Research*, 16: 447–465.
- Horgan, Terence and Timmons, Mark (1992a): 'Troubles on Moral Twin-Earth: Moral Queerness'. *Synthese*, 92: 450–474.
- Horgan, Terence and Timmons, Mark (1992b): 'Troubles for New Wave Moral Semantics: the 'Open Question Argument' Revived'. *Philosophical Papers*, 21: 153–175.
- Horwich, Paul (1998): *Meaning* (Oxford: Oxford University Press).
- Huemer, Michael (2005): *Ethical Intuitionism* (Houndmills: Palgrave MacMillan).
- Jackson, Frank (1998): *From Metaphysics to Ethics* (Oxford: Oxford University Press).
- Jackson, Frank and Pettit, Phillip (1995): 'Moral Functionalism and Moral Motivation'. *Philosophical Quarterly*, 45: 20–40.
- Jackson, Frank, Pettit, Philip, and Smith, Michael (2000): 'Ethical Particularism and Patterns'. In B. Hooker and M. Little (eds.): *Moral Particularism* (Oxford: Oxford University Press), 79–99.
- Kim, Jaegwon (1984): 'Concepts of Supervenience'. *Philosophy and Phenomenological Research* 45: 153–176.
- Kripke, Saul (1979): 'Speaker's Reference and Semantic Reference.' In P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language*. Minneapolis: University of Minneapolis Press.
- Kripke, Saul (1980): *Naming and Necessity* (Oxford: Blackwell).
- Kripke, Saul (1982): *Wittgenstein on Rules and Private Language* (Oxford: Blackwell).

- Laurence, Stephen, Margolis, Eric and Dawson, Angus (1999): 'Moral Realism and Twin Earth'. *Facta Philosophica*, 1: 135–165.
- Lenman, James (2006): 'Moral Naturalism'. Available at <<http://plato.stanford.edu/entries/naturalism-moral/>>
- Lewis, David (1972): 'Psychophysical and Theoretical Identifications'. *Australasian Journal of Philosophy*, 50: 249–258.
- Lewis, David (1974): 'Radical Interpretation'. *Synthese*, 27: 331–344.
- Lewis, David (1983): 'New Work for the Theory of Universals'. *Australasian Journal of Philosophy*, 61: 343–377.
- Lewis, David (1984): 'Putnam's Paradox'. *Australasian Journal of Philosophy*, 62: 221–236.
- Lewis, David (1986): *On Plurality of Worlds* (Oxford: Blackwell).
- Little, Margaret (1994): 'Moral Realism II: Non-Naturalism'. *Philosophical Books*, 35: 225–232.
- Loewer, Barry (1997): 'A Guide to Naturalizing Semantics'. In B. Hale and C. Wright (eds.): *A Companion to the Philosophy of Language* (Oxford: Blackwell), 108–126.
- Mackie, J.L. (1977): *Ethics – Inventing Right and Wrong* (London: Penguin Books).
- Majors, Brad (2005): 'Moral Discourse and Descriptive Properties'. *The Philosophical Quarterly*, 55: 475–494.
- McDowell, John (1998 [1981]): 'Non-Cognitivism and Rule-Following'. In his *Mind, Value, and Reality* (Cambridge, MA: Harvard University Press), 198–219.
- McPherson, Tristram (2012): 'Ethical Non-Naturalism and the Challenge of Metaphysical Supervenience'. *Oxford Studies in Metaethics*, 7: 205–234.

- Merli, David (2002): 'Return to Moral Twin Earth'. *Canadian Journal of Philosophy*, 32: 207–240.
- Merli, David (2009): 'Possessing Moral Concepts'. *Philosophia* 37: 535–556.
- Merrill, G.H. (1980): 'The Model-Theoretic Argument against Realism'. *Philosophy of Science*, 47: 69–81.
- Miller, Alexander (2007): *Philosophy of Language*, 2<sup>nd</sup> edition (London: Routledge).
- Millikan, Ruth (1984): *Language, Thought, and Other Biological Categories* (Cambridge, MA: MIT Press).
- Moore, G.E. (1903): *Principia Ethica* (New York: Cambridge University Press).
- Nagel, Thomas (1996): 'Universality and the Reflective Self'. In O. O'Neill (ed.), *Christine Korsgaard, The Sources of Normativity* (Cambridge: Cambridge University Press), 200–209.
- Papineau, David (1979): *Theory and Meaning* (Oxford: Oxford University Press).
- Parfit, Derek (2011): *On What Matters*, Vols. 1–2 (Oxford: Oxford University Press).
- Peacocke, Christopher (1987): 'Understanding Logical Constants: A Realist's Account'. *Proceedings of the British Academy*, 63: 153–199.
- Peacocke, Christopher (1992): *A Study of Concepts* (Cambridge, MA: MIT Press).
- Putnam, Hilary (1975): 'The Meaning of 'Meaning''. *Minnesota Studies in the Philosophy of Science*, 7: 221–274.
- Putnam, Hilary (1977): 'Realism and Reason'. *Proceedings of the American Philosophical Association*, 50: 483–498.
- Putnam, Hilary (1981): *Reason, Truth, and History* (Cambridge: Cambridge University Press).
- Quine, W. V. (1960): *Word and Object*. Cambridge, MA: MIT Press.

- Ross, W.D. (2002 [1930]): *The Right and the Good*, reprinted with an introduction by P. Stratton-Lake (Oxford: Oxford University Press).
- Russell, Bertrand (1905): 'On Denoting'. *Mind*, 14: 479–493.
- Scanlon, T.M. (2014): *Being Realistic about Reasons* (Oxford: Oxford University Press).
- Schroeder, Mark (2007): *The Slaves of Passions* (Oxford: Oxford University Press).
- Schroeter, Laura and Schroeter, François (2003): 'A Slim Semantics for Thin Moral Terms'. *Australasian Journal of Philosophy*, 81: 191–207.
- Schroeter, Laura and Schroeter, François (2007): 'Normative Realism: Co-Reference without Convergence'. *Philosophers' Imprint*, 13.13: 1–22.
- Schwarz, Wolfgang (2014): 'Against Magnetism'. *Australasian Journal of Philosophy*, 92: 17–36.
- Searle, John (1958): 'Proper Names'. *Mind*, 67: 166–173.
- Setiya, Kieran (2011): 'Review of Derek Parfit's *On What Matters*'. *Mind* 120: 1281–1288.
- Shafer-Landau, Russ (2003): *Moral Realism – a Defence* (Oxford: Oxford University Press).
- Shafer-Landau, Russ (2006): 'Ethics as Philosophy: A Defence of Ethical Nonnaturalism'. In R. Shafer-Landau (ed.): *Metaethics after Moore* (Oxford: Oxford University Press), 209–232.
- Sider, Theodore (2011): *Writing the Book of the World* (Oxford: Oxford University Press).
- Smith, Michael (1994): *The Moral Problem* (Oxford: Blackwell).
- Strawson, Peter (1950): 'On Referring'. *Mind*, 59: 320–344.
- Suikkanen, Jussi (2010): 'Non-Naturalism: The Jackson Challenge'. *Oxford Studies in Metaethics*, 5: 87–110.

- Suikkanen, Jussi (2016): 'Non-Realist Cognitivism, Truth and Objectivity'. Forthcoming in *Acta Analytica*.
- Sundell, Timothy (2012): 'Disagreement, Error, and an Alternative to Reference Magnetism'. *Australasian Journal of Philosophy*, 90: 743–759.
- van Roojen, Mark (2006): 'Knowing Enough to Disagree'. *Oxford Studies in Metaethics*, 1: 161–194.
- van Roojen, Mark (2015): *Metaethics – A Contemporary Introduction* (London: Routledge).
- Weatherson, Brian (2003): 'What Good Are Counterexamples?'. *Philosophical Studies*, 115: 1–31.
- Wedgwood, Ralph (2001): 'Conceptual Role Semantics for Moral Terms'. *The Philosophical Review*, 110: 1–30.
- Wedgwood, Ralph (2007): *The Nature of Normativity* (Oxford: Oxford University Press).
- Wielenberg, Erik (2014): *Robust Ethics – the Metaphysics of Godless Normative Realism* (Oxford: Oxford University Press).
- Williams, Robert (2007): 'Eligibility and Inscrutability'. *Philosophical Review*, 116: 361–399.
- Williams, Robert (2015): 'Lewis on Reference and Eligibility'. In B. Loewer and J. Schaffer (eds.): *A Companion to David Lewis* (Oxford: Wiley-Blackwell), 367–381.
- Williamson, Timothy (2001): 'Ethics, Supervenience, and Ramsey Sentences'. *Philosophy and Phenomenological Research*, 62: 625–630.