

Rescuing tracking theories of morality

Marc Artiga¹

Published online: 1 April 2015
© Springer Science+Business Media Dordrecht 2015

Abstract Street’s (Philos Stud 127(1):109–166, 2006) Darwinian Dilemma purports to show that evolutionary considerations are in tension with realist theories of value, which include moral realism. According to this argument, moral realism can only be defended by assuming an implausible tracking relation between moral attitudes and moral facts. In this essay, I argue that this tracking relation is not as implausible as most people have assumed by showing that the three main objections against it are flawed. Since this is a key premise in the reasoning, I conclude that the Darwinian Dilemma against moral realism can be resisted.

Keywords Moral realism · Darwinian Dilemma · Teleosemantics · Debunking arguments · Sharon Street

1 Introduction

A significant amount of current metaethical research is devoted to investigate the consequences of assuming that moral cognition is the product of evolutionary forces. Evolutionary theories of morality typically hold that natural selection was the most important process in shaping our tendency to produce moral beliefs.¹ More precisely, they usually hold that moral cognition is an adaptation, i.e. that our tendency to produce moral judgments is the product of natural selection (Joyce

¹ I will assume throughout the paper that moral cognition involves moral beliefs which are truth-apt, since probably only cognitivist views can employ the debunking argument discussed here (Mason 2010).

✉ Marc Artiga
marc.artiga@gmail.com

¹ Universidad Nacional Autónoma de México- Instituto de Investigaciones Filosóficas, Circuito Maestro Mario de la Cueva, s/n. Ciudad Universitaria, Coyoacán, 04510 Mexico, D.F., Mexico

2006, ch. 4).² The exact details of this process are still disputed (Fraser 2010; Joyce 2013; Sterelny et al. 2013), but at least there seem to be good reasons for thinking that moral cognition has been strongly shaped by selective pressures (Street 2006, pp. 115–121).

Additionally, many think that the idea that moral cognition is an adaptation has groundbreaking consequences concerning the existence of moral facts. In particular, it has been argued that this claim, together with some plausible empirical assumptions, entails (or, at least, strongly suggests) the falsity of traditional forms of moral realism (Joyce 2001a, 2006; Ruse 1996; Street 2006). In this paper I would like to critically examine one way of developing this argument.

Note that, *prima facie*, the claim that a cognitive mechanism is an adaptation does not imply or suggest that it is an unreliable system, as evidenced by the fact that our perceptual mechanisms are adaptations and nonetheless highly reliable (Fraser 2010; Wilkins and Griffiths 2013). Thus, a crucial question is what kinds of conditions and evolutionary processes must be in place for a reliable mechanisms to evolve and whether these conditions obtained in the case of moral cognition. In the context of our discussion, this is usually spell out in terms of a tracking relation between moral beliefs and moral facts. The issue, then, is whether the evolutionary process that shaped moral cognition established a tracking or a non-tracking relation between moral beliefs and moral facts. With respect to that question, there seems to be a wide consensus among metaethicists on the rejection of the tracking thesis as a viable approach to moral cognition (Blackburn 1993, p. 168; Copp 2008; Fraser 2014; Gibbard 1990; Joyce 2006; Kahane 2011; Ruse 2005; Street 2006; Wilkins and Griffiths 2013). A central goal of this essay is to show that this consensus is ill-founded.

To defend the plausibility of the tracking thesis, I will primarily focus on Street's (2006) Darwinian Dilemma and her discussion, since it is probably the most clear defense of the idea that moral beliefs are off-track.³ Thus, I will first briefly outline the Darwinian Dilemma, and I will concentrate on the alleged implausibility of the tracking thesis concerning moral beliefs. Afterwards, I will spell out in some detail the three most common considerations against it and I will show that all of them are unconvincing. If I am right, then, the consensus about the inadequacy of the tracking thesis is unfounded. The tracking theory of morality can be rescued and moral realism might be able to defeat this influential objection.

1.1 The dilemma

Now, a preliminary difficulty in addressing this objection is that there is no general agreement on the structure of argument (see, for instance, Clarke-Doane 2012;

² The claim that moral cognition is an adaptation does not entail that particular moral beliefs have been selected for. One might think that selection has acted on some general moral rules or principles (Hauser 2006), on a capacity for moral concepts (Joyce 2006) or on some indirect mechanism that tends to produce moral attitudes (Street 2006; Joyce 2013, p. 558). Similarly, the claim that the moral sense is an adaptation does not entail that there is some dedicated cognitive system (Fraser 2014; Joyce 2013).

³ Street's argument is not only directed against moral realism, but against all realist theories of value. Here I will exclusively focus on the case of morality.

Enoch 2010; Kahane 2011; Schafer 2010; Shafer-Landau 2012). Furthermore, as I will show, how the dilemma is cashed out has important consequences for some of the key arguments in the discussion. Thus, let us start by formulating Street's objection.

As I said, the Darwinian Dilemma is a debunking argument, whose main goal is to undermine moral realism. Of course, the exact definition of moral realism is highly controversial (Kahane 2011), but the following rough approximation will do for our purposes:

MORAL REALISM: There are moral facts, which are independent of all our (actual or ideal) moral attitudes. (Street 2006, p. 110)

According to MORAL REALISM, there are indeed moral facts and they do not constitutively depend on the attitudes we are inclined to take. More precisely, moral facts are neither determined by our actual evaluative judgments, nor by the judgments we would make under certain idealized conditions.⁴

Debunkers try to show that MORAL REALISM clashes with a set of very plausible hypotheses. For this reason, their strategy can be usefully expressed in terms of a set of mutually inconsistent claims. In particular, Street (2006; and also Joyce 2006) argue that these three theses are incompatible:

(A) MORAL REALISM

(B) Evolutionary forces have played a tremendous role in shaping our tendency to hold certain moral beliefs.

(C) Some of our moral beliefs are justified

Street's (2006) Darwinian Dilemma is supposed to show that these three theses are inconsistent by arguing that (A) and (B) entail the denial of (C) (Vavova 2015, p. 107). Given the plausibility of (B) and (C), she concludes that (A) should be rejected. The goal of this essay is to show that (A), (B) and (C) are fully compatible and that, as a result, we lack convincing reasons for abandoning any of them.

Let us now focus on Street's argument. As we saw, the first assumption is that evolutionary forces had a enormous influence in devising our moral attitudes (i.e. B). This is of course an empirical thesis, so this claim is subject to the usual standards of scientific research (FitzPatrick 2015), but there seem to be significant evidence in its favor (Joyce 2006, ch. 1; Street 2006, pp. 115–121). In any case, if one is not moved by these considerations, debunking arguments should simply be conditionalized: they show what would be the case *if* this empirical thesis were true. For the sake of the discussion, I will take this claim for granted here.

The second premise is obviously MORAL REALISM, i.e. the claim that there are moral facts, which are independent from our (actual or ideal) evaluative attitudes. The Darwinian Dilemma is supposed to show that these two premises entail that all our moral judgments are unjustified.

⁴ Street (2006, p. 138) adds extra requirements, but since they are irrelevant for our discussion, I suggest to leave them aside.

Once it is assumed that there are independent moral facts and that evolutionary forces have strongly shaped the cognitive capacities that led us to hold certain moral beliefs, the next step is to question the relation between them. Here is where Street's (2006) dilemma comes in: suppose there has been no relation whatsoever between moral truths and the evolution of moral cognition; that result may potentially undermine our confidence in moral beliefs. After all, if in the process explaining our moral attitudes did not intervene moral facts, the mere observation that we accept certain moral beliefs constitutes no evidence in favor or against their truth. In the same way that discovering that your belief that Proust wrote *In Search of Lost Time* was caused by the ingestion of a pill would undermine your credence on Proust's work, if the etiology of moral cognition has no relation whatsoever with moral truths, moral judgments would be unjustified (Joyce 2006, p. 179).

But suppose we take the other horn of the dilemma. Let us imagine for a moment that there is some relation between moral truths and moral beliefs. According to Street, the only relation that could support our confidence in moral beliefs is a tracking relation. In the same way that discovering the evolutionary origin of perceptual mechanisms does not undermine the justification of perceptual experiences because we think it has established a tracking relation, if there has been such a relation between moral beliefs and moral facts, the etiology of our beliefs would not jeopardize their reliability.

Unfortunately, Street and many others maintain that the non-tracking hypothesis is by far the most plausible one. Thus, moral attitudes are off-track and, hence, unjustified. This is basically the structure of Darwinian Dilemma, which can be spelled out as follows:

Darwinian Dilemma (DD):

1. Evolutionary forces have played a tremendous role in shaping out tendency to hold certain moral beliefs.
2. MORAL REALISM is true.
3. Given 1 and 2, either (a) there has been a relation between the evolutionary forces shaping our tendency to produce moral beliefs and independent moral truths or (b) there has not been such a relation.
4. If there has been no relation, then evolutionary forces have been a purely distorting influence (in which case our moral beliefs would be unjustified).
5. If there has been a relation, only a tracking relation can avoid the conclusion that evolutionary forces have been purely distorting (in which case our moral beliefs would be unjustified).
6. The relation between evolutionary forces and independent evaluative truths has not been a tracking relation.

(.∴) Our moral beliefs are unjustified.

In this argument, the first two main theses are (A) and (B) and the conclusion is the denial of (C). So, if the DD holds it would vindicate the incompatibility of (A), (B) and (C). Interestingly, note that the realist can not simply bit the bullet and accept the denial of (C). For, suppose our moral beliefs are unjustified. Since probably our moral beliefs are the best evidence we have for the existence of certain

moral facts, accepting that or moral beliefs are unjustified would threaten to undermine MORAL REALISM. So the rejection of (C) seems to lead quite straightforwardly to a powerful argument against (A).

Alternatively, the realist could attempt to deny that evolutionary forces have played a enormous role in shaping our tendency to hold certain moral beliefs, but as we saw, there seems to be a significant body of evidence in its favor. Moreover, denying that moral cognition is an adaptation would probably make things even harder for the realist (Street 2006, p. 142). As a consequence, debunkers suggest we should abandon MORAL REALISM.

Given this powerful argument, there are not many options open for the realist. On the one hand, the moral realist can not sensibly deny 1 or 2 (that is, A and B) and, on the other, 3–5 seem to rightly describe the space of possible options. Yet 6 is also a substantive empirical thesis required for the argument to go through. This is the premise I will take issue with. In particular, I will argue that the reasons given in favor of this premise do not withstand serious scrutiny. If I am right, the tracking thesis is a viable hypothesis that needs to be seriously taken into account. Consequently, unless additional arguments are provided in its favor, premise 6 is unjustified and the Darwinian Dilemma can be resisted.

Now, since my goal is to consider whether the relation between evolutionary forces and moral truths have been tracking or non-tracking, we need first be more clear about these hypotheses.⁵ First of all, following Copp (2008, p. 194), let us call the claim denied in 6 the ‘Tracking Thesis’:

Tracking Thesis (TT): Darwinian forces caused our moral beliefs to track moral facts.

According to the DD, the Tracking Thesis (TT) is the claim needed by the moral realist. It states that evolutionary forces have shaped the mechanisms responsible for our tendency to hold certain moral beliefs in such a way that these beliefs tend to correspond to real moral facts. But what kind of evidence could be provided in favor or against this thesis? Street argues that TT requires the truth of the Tracking Account:

Tracking Account (TA): Tendencies to make certain kinds of moral judgments rather than others contributed to our ancestor’s reproductive success because they constituted true representations of moral facts.

Street maintains that TA is the only process that could vindicate TT.⁶ Furthermore, note that this explanation is quite demanding. It claims that our tendencies to hold certain moral beliefs exist in part because they were true. Thus, MORAL REALISM could be secured only if one could show that the fact that moral beliefs were true

⁵ I would like to thank an anonymous referee for helpful comments on this section.

⁶ Of course, one can come up with other explanations that would vindicate TT. For instance, that moral cognition is the product of a benevolent God that causes us to hold only true moral beliefs. Nonetheless, it is generally granted that TA is the only plausible hypothesis that is in accordance with a broad scientific perspective and with the assumption that evolutionary forces have played a tremendous role in shaping our tendency to hold moral beliefs.

played some important role in the explanation of our tendency to accept determinate moral propositions.

Yet there are many alternative hypotheses, which entail that our moral beliefs are off-track, i.e. that TT is false (see Clarke-Doane 2012, p. 325). The idea that our tendencies to hold certain moral beliefs are the product of genetic drift, for instance, would probably lend to support to the off-track hypothesis. As a general rule, Street suggests that the process giving rise to moral cognition has probably been non-tracking if the truth of moral beliefs played no role in the explanation of our tendency to accept certain moral propositions.⁷ Put in a different way, if there has been no relation between the evolution of moral cognition and the moral truths, moral beliefs are off-track and hence unjustified.⁸ If the evolution of moral cognition has been unrelated to the existence of moral truths, our moral beliefs should be distrusted.

With this idea in mind, Street puts forward the following alternative hypothesis:

Output Account (OA): Tendencies to make certain kinds of moral judgments rather than others contributed to our ancestor's reproductive success because they forged adaptive links between our ancestors' circumstances and their responses to those circumstances, getting them to act, feel and believe in ways that turned out to be reproductively advantageous. (based on Street 2006, p. 127)

Street suggests that the OA would undermine the TT. According to OA, moral beliefs were selected for because they led our ancestors to behave in fitness-enhancing ways. In contrast, according to TA what explains the selection of moral cognition is the fact that they allowed our ancestors to truly represent moral facts. TT (and, hence, premise 6 of the Darwinian Dilemma) depends on the whether the TA or the OA are true.

To illustrate the contrast between an explanation in terms of TA and in terms of OA, take the belief that the fact that something would promote the interests of a family member is a reason to do it. If TA is right, the selection of the tendency to produce this belief was favored by natural selection because it helped our ancestors to be aware of a moral fact. In contrast, OA would maintain that thinking in these terms caused us to be more prone to behave in ways that were beneficial for us, our

⁷ Street defines a non-tracking relation as involving an independence of a belief from its truth *or falsity*. That would suggest that a tracking relation should be defined as the dependence of a belief on its truth *or falsity*. Yet there can be no tracking relation if moral beliefs are false. Furthermore, moral realism is one of the premises of the argument, so the tracking and non-tracking accounts should be defined by reference to the independent moral facts. For these reason, I exclusively focus on the relation between beliefs and moral truths.

⁸ Again, for the process to be off-track one need not assume that truths are completely independent of our beliefs. It suffices if the relation is of the wrong kind. If, for instance, I believe that neuroscientists are evil without having any evidence for it, but merely because an evil neuroscientist has directly stimulated my brain, my belief and its truth would not be independent, but my belief would surely be unjustified. Nonetheless, since these cases are fairly uncommon (and very implausible in the case of morality) I will leave them aside.

kin or our group. On this view, the existence of moral facts played no role in the explanation of moral cognition.

Many take OA to be much more plausible than TA. And since TA is regarded as the only way of defending TT, they conclude that TT should be abandoned. In the remainder of the paper, I would critically examine the arguments in support of this claim. Let me hasten to add, however, that I will not attempt to show that TT or TA are true; after all, that would require a defense of MORAL REALISM, which lies beyond the scope of this essay. The more modest goal (but, I hope, still interesting) is to argue that we lack convincing reasons against TT, because all the objections against TA fail. If that is right, then, premise 6 is far from settled, in which case the moral realist should not be worried by the Darwinian Dilemma.

2 Defending the tracking account

In this section I will discuss the three main arguments purporting to show that the OA is preferable to TA [with special emphasis on Street (2006)] and I will argue they are all unconvincing. As a result, I will conclude that premise 6 has not been established and that, unless additional arguments are provided, the DD can be resisted.

2.1 Parsimony argument

The first objection against TA appeals to Occam's Razor. It is part of common wisdom that, other things being equal, if a theory T_1 is more ontologically parsimonious than a theory T_2 , it is rational to prefer T_1 to T_2 . Following this dictum, many argue that OA is preferable to TA because it is more ontologically parsimonious. Both theories can explain why we are endowed with a tendency to hold certain moral beliefs, but TA posits independent moral truths, while the OA dispenses with them. OA does not presuppose the existence of evaluative facts because it explains how we evolved moral attitudes by appealing to the behaviors that those attitudes caused and their promoting reproductive success. Consequently, given that both have the same explanatory power but one of them postulates less entities (in particular, moral facts), it is rational to prefer OA in terms of ontological economy (Gibbard 1990; Harman 1977; Joyce 2006, p. 188; pp. 107–108; Ruse 2005, p. 176; Schafer 2010; Sinclair 2012, p. 652; Street 2006, p. 129).

I think there are at least three reasons why this argument is unconvincing. First of all, note that if TA is less parsimonious, it is in virtue of assuming the existence of moral facts and not in virtue of assuming a tracking relation. To put it in a slightly different way: there is nothing in TA besides its commitment to moral facts that is less parsimonious than OA. However, recall that the Darwinian Dilemma assumes MORAL REALISM as a fundamental thesis (Kahane 2011; Shafer-Landau 2012; Vavova 2015). In the formulation given above, the key question is whether (A) and (B) entail that all our moral judgments are unjustified, so MORAL REALISM is a central premise in the argument (see footnote 7). Therefore, since in any case the existence of moral facts is assumed when assessing the parsimony of TA and OA, TA is not

less parsimonious than OA. If MORAL REALISM is taken for granted, Occam's Razor cuts no ice.

Secondly, even if one thinks premise 2 should not be taken into account when assessing the parsimony of TA and OA, a further difficulty with the argument is that it fails to provide any additional worry for the realist. Certainly, we already knew before the Darwinian Dilemma was raised that MORAL REALISM might be less parsimonious than moral anti-realism,⁹ so if one rejects MORAL REALISM only for this reason, it fails to provide any new challenge for the realist. Let me illustrate the idea with the following example: suppose two arguments are raised against the existence of angels. The first one appeals to the fact that a world with angels is less parsimonious than a world without angels. The second argument claims that a world in which our beliefs on angels track angels is less parsimonious than a world in which our beliefs on angels do not track anything. We would unhesitatingly object that the second argument does not add anything substantial to the first one. Similarly, if one uses parsimony considerations against realism and those are the only reasons in favor of a key premise of DD, it fails to provide new evidence against the realist.

Finally, there are already some theories that purport to show that moral facts do not constitute a substantial addition to one's ontology. In particular, it is not obvious to what extent naturalist theories of morality, which assert that moral facts are constituted (or reducible) to non-moral facts, are less parsimonious than eliminativist theories (Copp 2008, p. 190; Shafer-Landau 2003, p. 114). For instance, if moral facts reduce the certain social facts (whose existence anyone is willing to accept anyway) a tracking theory holding that evolutionary forces led us to track these social facts would not be less parsimonious than a theory that presupposes them but explains moral cognition in different terms.¹⁰

Therefore, I doubt parsimony considerations can help decide between TA and OA.

2.2 Causal argument

The second argument is supposed to point out a difficulty of TA, rather than directly showing a virtue of OA. Some philosophers question the plausibility of TA because they doubt that the fact that moral beliefs are true can provide any evolutionary advantage at all (Joyce 2006, p. 183; Sinclair 2012, p. 652; Street 2006, pp. 129–131). In particular, Street (2006, pp. 132–133) writes:

The [Output Account] is much clearer than the tracking account, which turns out to be rather obscure upon closer examination. As we have seen, according

⁹ An interesting question is whether Street's constructivism is more parsimonious than Moral Realism.

¹⁰ Still, one could press this objection further and argue that many of the non-moral facts postulated by reductionists are not posited by the supporter of the OA, so there is some gain in ontological parsimony. The problem of this rejoinder is that at this point it becomes much less clear whether OA is actually more parsimonious than TA. Once it is granted that moral facts can be reduced to non-moral facts, does TA (which explains our tendencies to produce moral beliefs by appealing to certain non-moral facts) postulate less entities than OA (which requires moral beliefs to have certain effects)?

to the tracking account, making certain evaluative judgments rather than others promoted reproductive success *because these judgments were true*. Bet let's now look at this. How exactly is that supposed to work? (...) A creature obviously can't run into [moral facts] or fall over them or be eaten by them. In what way would it have promoted the reproductive success of your ancestors to grasp them?

In a nutshell, the question is why would it promote an organism's reproductive success to truly represent independent moral facts. Of course, in general having true beliefs is better than having false beliefs or not having any beliefs at all (Wilkins and Griffiths 2013). For instance, in the long run it is obviously better to have true perceptual beliefs than having them false. However, moral attitudes seem to be different in that respect; it is unclear why the truth of moral claims could make any difference concerning survival and reproduction.

I think this worry stems from questioning whether moral facts can figure in causal explanations.¹¹ It is not obvious, for example, that the wrongness of setting a wood on fire has any causal influence. On this reading, the issue at stake is whether the presence or absence of moral facts can make a difference concerning survival and reproduction. If they are causally inert, the evolutionary advantage of truly tracking these entities (over merely *thinking* that one is tracking them) is unclear. Entities that do not figure in causal explanations cannot play a role in evolutionary explanations. If that interpretation is on the right track, the relevant issue is whether moral facts are causally efficient. Intuitive doubts concerning the causal powers of moral properties diminish the plausibility of the TA.

I believe this objection points at substantive question that realists should address. However, two of the previous replies seem to apply here as well. First, we already knew that the moral realist needs to account for the causal efficacy of moral facts (for instance, in order to account for moral knowledge). So, if the Darwinian Dilemma exclusively relies on this idea, it fails to provide any additional argument against MORAL REALISM. Consider again a philosopher intending to provide two arguments against the existence of angels. The first one appeals to the fact that it is not obvious how angels can be causally efficient. The second argument claims that probably our beliefs on angels cannot track angels, because it is not clear that angels figure in causal explanations. Again, the second argument is completely parasitic on the first one and fails to provide any additional reason for denying the existence of angels.¹²

Secondly, it should be noticed that there are already some replies available in the literature. For instance, if one holds a (type or token) physicalist theory of morality according to which moral facts are constituted by natural facts and also thinks that the constituting natural facts are not epiphenomenal, then one can plausibly argue

¹¹ Alternatively, this intuition of obscurity might be rooted in the queerness of moral facts (Mackie 1977). However, I doubt Street or Joyce are grounding their criticisms on that claim (and, in any case, I provide a response for that worry below).

¹² The same argument, *mutatis mutandis*, would show that the queerness of moral facts should not be used against TA (see 11).

that moral facts are not epiphenomenal either (Brink 1989, ch. 8).¹³ Similarly, some non-naturalists like Shafer-Landau think that moral properties cannot be reduced to non-moral properties, and nevertheless maintain that they inherit their causal profile from the non-moral properties that constitute them.¹⁴ So there are different ways of showing that moral properties can have causal effects, which would provide an answer to this worry (see Joyce 2006, p. 184).¹⁵

Street (2006, p. 131) seems to consider an answer along these lines and replies:

At least so far, this is not much of an explanation, either. What kinds of natural facts are we talking about and exactly why did it promote reproductive success to track them? The naturalist can certainly try to develop answers to these questions, but at least on the face of things, the prospects appear dim.

If any of strategies I considered succeeds in showing that moral facts are causally relevant, their explanatory value could be secured. Crucially, note that my argument does not have to assume any specific answer to this worry. Rather, my point is that, since Street's view is a particular instance of a more general concern in which there is already a substantive body of literature, her argument depends on a very controversial assumption. Accordingly, for Street's reply to be compelling, she should have to show that all strategies purporting to defend moral facts can figure in causal explanations are unsuccessful.¹⁶ Relying the defense of OA on this disputed and complex issue is, to say the least, a very unstable position.¹⁷

2.3 The abductive argument

In my opinion, the last objection is the most powerful and the one that has convinced more people that a Tracking Account is doomed to failure (Blackburn 1993, p. 168; Fraser 2014; Joyce 2006, p. 215; Kitcher 2005, p. 176; Street 2006, p. 132; Wilkins and Griffiths 2013). It will also be more complex to show why it is wrong.

The argument is apparently very simple, but it is backed by a very strong intuition. Suppose it is granted that moral facts can figure in causal explanations and

¹³ Note that Street (2006, p. 145) explicitly claims her dilemma also undermines these forms of naturalism.

¹⁴ If one holds a non-reductionist theory, there might be overdetermination problems (Kim 1998). However, this difficulty also affects many others kinds of facts, like mental or biological facts. Thus, if one denied the causal efficacy of moral properties by appealing to these overdetermination problems, one would probably have to deny that all these entities can figure in causal explanations (Shafer-Landau 2003, ch. 4). Accordingly, one would be forced to deny a Tracking Account of mental or biological facts and *this* is very implausible.

¹⁵ Even Enoch (2010), who endorses robust moral realism but rejects the causal efficacy of moral facts, thinks that there is a grounded correlation between moral beliefs and moral facts, such that they can make a difference concerning survival and reproduction.

¹⁶ And again, it is not obvious that Street's (2006) own theory does not fall prey to this problem.

¹⁷ Indeed, in my response to the third objection I will explain a bit more in which sense causally relevant moral facts could play an important explanatory role in the evolution of moral cognition.

that TA is not less parsimonious than OA. Still, one can argue for the superiority of OA with an inference to the best explanation:

ABDUCTION: The Output Account offers a better explanation than the Tracking Account of the fact that we have certain moral attitudes rather than others.

Consider a particular case: Why is the tendency to belief that we have special obligations towards our children so widespread? The answer offered by OA is that those ancestors who made these judgments tended to act more kindly towards their children, and for this reason they did better than those ancestors lacking these beliefs. More generally, having beliefs concerning one's special duties towards Ps increases the probability of one's behavior benefiting Ps. So, under the assumption that Ps' fitness is somehow related to one's own fitness (e.g. by kin), the effects of this behavior will explain why this tendency is selected for.

In contrast, consider the slim explanatory import of TA: it merely claims that the belief that we have special obligations towards our children is actually so widespread in humans because we actually do have these obligations. That is, the only explanation provided by the TA appeals to the truth of the belief. Most people think OA does a much better job in explaining the current existence of moral beliefs.

Street's contention is even stronger: she doubts TA offers an explanation at all. And, indeed, I think she is partially right. TA alone does not provide a satisfactory explanation of our tendencies to produce moral judgments. Nevertheless, in what follows I will argue that we have independent reasons for thinking that (1) OA is compatible with TA (and with TT) and that (2) in any case, it is unfair to assess the explanatory value of TA alone (i.e., without OA). To make these two points, however, will require a brief consideration of evolutionary theories of representation. That will complete my defense of the Tracking Thesis.

2.3.1 *Naturalistic theories of representation*

TA and the OA disagree on the process that accounts for the selection of our tendencies to produce moral beliefs. According to the former, tendencies to produce certain moral beliefs were selected because they allowed our ancestors to track moral truths, while the latter merely appeals to their effects on behavior. As I said, most people think the second view offers the best explanation. But is there some way of settling this issue without relying on our intuitions in the moral case? Can we provide some independent reasons for favoring any of these options? I think we can. This is one of the reasons I suggest to consider naturalistic theories of representation.

Arguably, moral beliefs are certain kind of representational states.¹⁸ So the question of why moral beliefs evolved is a particular case of the more general question of why certain representational systems evolve. Fortunately, there are already some theories that try to specify which conditions must be in place for a

¹⁸ Recall that in this discussion both TA and OA assume cognitivism, i.e., the claim that moral attitudes involve truth-apt moral beliefs. TA is committed to this claim because it holds moral attitudes represent moral facts, and OA because it is used as a premise in an argument whose conclusion is that *moral beliefs* are off-track (see Mason 2010).

representational system to arise. Since here it is assumed that moral cognition is an evolved representational system, I think there are interesting lessons to be learned from this debate.

Paying attention to evolutionary theories of representation have several interesting advantages. On the one hand, it provides a strategy for addressing the debate between the OA and TA that is independent from the metaethical discussion. Furthermore, since most current theories of representation agree on certain key ideas, progress is to be reasonably expected. Finally, a popular strand among theories of representation, namely teleological theories, appeal to evolutionary considerations in order to explain the existence and content of representational mechanisms. In that respect, it should be obvious why a (cognitivist) evolutionary theory of morality should look closely at evolutionary theories of mental representation. We will see that this intuition is on the right track.¹⁹

More precisely, *Teleological* or *Teleosemantic Theories* of representation are naturalistic approaches (Artiga 2014; Godfrey-Smith 1996; Martinez 2013; Millikan 1984; Neander 1995, 2013; Papineau 1987; Price 2001; Schulte 2015; Shea 2007). Their main goal is to describe the conditions and processes that give rise to representational mechanisms without appealing to further representational systems. Of course, there are some disagreements between different versions, but for simplicity I will describe a basic framework that I think could be accepted by most teleosemanticists.²⁰

The most common teleosemantic framework strongly relies on two notions: etiological function and sender-receiver structure (Godfrey-Smith 1996; Millikan 1984). On the one hand, according to the etiological notion of function, functions are selected effects (Neander 1991). More precisely, functions are those effects of a trait that explain why it was selected for by natural selection. For instance, the pancreas' function is to produce hormones (including insulin, glucagon, somatostatin) and pancreatic juice because these are the effects that explain why organisms having pancreas were favored by natural selection. Secondly, teleological theories appeal to sender-receiver systems, which are mechanisms composed of two sub-systems: a sender, which produces a set of states given certain inputs, and a receiver, which acts in certain ways given the states produced by the sender.

Now, teleosemantics combines the notions of etiological function and sender-receiver structure in order to provide a naturalistic account of representation.²¹ A representational system is a mechanism involving a sender-receiver structure, in which sender and receiver are endowed with the right etiological functions. More precisely, the function of the sender is to produce a state R (the *representation*)

¹⁹ Although the success of teleological theories in naturalizing intentionality is a hotly disputed issue, there are at least good reasons for taking these theories seriously: they seem to yield the right results in a wide range of cases, can account of misrepresentation and provide a fully naturalistic account of content (Gibbard 1990, ch. 6; Harms 2000; Neander 2012; Sinclair 2012).

Nonetheless, it is important to keep in mind that my argument does not require teleosemantics to be true. I merely use it in order to show what is wrong in the reasoning leading to the abductive argument.

²⁰ This framework was first developed by Millikan (1984).

²¹ Some teleosemanticists do without senders (Cao 2012; Stegmann 2009) or without receivers (Dretske 1995; Neander 1995; Schulte 2015). Nonetheless, I think all of them could accept that in most cases both structures are in place. That suffices for my argument to go through.

when another state *S* obtains (the *representatum*), while the function of the receiver is to produce an effect when state *R* is produced. If that structure obtains, then *R* is a representation and means *S*.²²

Let me illustrate this theory with an example. Cockroaches and crickets have two short appendages that extend from the rear of their abdomen called ‘cerci’. Each cercus has a set of ‘slender filiform sensory hairs’ (Comer and Leung 2004, p. 314), which are sensitive to air movements. Each hair is associated with an efferent neuron, such that when a particular hair senses air moving at a certain velocity (which usually enough corresponds with the presence of a predator), cockroaches respond with an evasive behavior in the opposite direction. Needless to say, this mechanism has helped cockroaches and crickets to avoid being devoured by predators. This is a simple case in which the teleosemantic framework easily applies: the sender is the efferent neuron, the receiver is the motor system that generates the escaping behavior and the activation of the efferent neuron constitutes the representation that means something like *there is a predator in such and such direction*.

Notice that merely producing states that correlate with a predator being around does not explain why the representational system was selected. Teleosemantics teaches us that the representational system exists because (1) there was a *sender* that often enough produced neuronal activations when there was a predator around and (2) there was a *receiver* that acted in certain ways when neurons were activated. Both the truth of the representation and the adaptiveness of the behavior must be mentioned in an explanation of how the representational system evolved; neither of these conditions alone suffices. This is enough, I think, for showing what is wrong with Street’s abductive argument.

2.3.2 A reply to the abductive argument

From the discussion of naturalistic theories of representation, two lessons follow. The first one is that teleosemantics illustrates how OA and TA (and TT) can be true at the same time. The second argument suggests that in general one should not assess the explanatory value of TA alone. These two replies show why the abductive argument probably fails.

Let us begin with the compatibility argument. There are two possible ways of developing it²³: either by arguing that OA is compatible with TA (which entails TT), or by directly showing that OA is compatible with TT. I will present each of these strategies and show that they lead to the same result.

Consider first OA. According to this account, we have the tendency to produce certain moral beliefs because ‘they contributed to our ancestor’s reproductive success by forging adaptive links between our ancestors’ circumstances and their responses to those circumstances, getting them to act, feel and believe in ways that

²² Senders and receivers can be different organisms or different systems within the same organism. In the present discussion, we are mostly interested in the latter.

²³ I would like to thank an anonymous referee for helping me clarify the arguments in this section.

turned out to be reproductively advantageous' (Street 2006, p. 129). That is in complete accordance with the teleosemantic theory: any representational system contains a receiver and the selection of the representational system is partially explained by the fact that it acted in fitness-enhancing ways. So, from the assumption that moral cognition is an adaptation involving truth-apt representational states, it follows that part of the explanation of why we have this mechanism appeals to the fact that moral beliefs led our ancestors to behave in certain ways. Teleosemantics provides a robust and independent justification for the truth of OA in the context of moral judgments.

However, consider now TA. What is the reason for denying that the truth of moral beliefs partially explains the selection of the representational mechanism? Since both the parsimony and the causal argument have been blocked, the abductive argument can only suggest that OA offers a better explanation than TA. However, one explanation is better than an other only when they are alternatives. This is precisely what is denied by teleosemantic theories. Teleosemantics claims that the existence of a representational mechanism requires a sender (which often enough must represent truly) and a receiver (which often enough must lead to fitness-enhancing behavior). Both the truth of the representation and the adaptiveness of the ensuing behavior explain why a representational mechanism exists at all. Thus, TA and OA should not be considered alternative accounts, but complementary explanations. Consequently, the fact that OA is true does not diminish at all the plausibility of TA.

A caveat is important here. Note that teleosemantics does not vindicate TA. Among other things, to show that TA is true one would need to argue that there are moral facts. What teleosemantics is supposed to illustrate is that we lack convincing reasons against TA; in particular, that even if OA were true, that would do nothing to show that TA is false. As a result, the only evidence in favor of premise 6 vanishes and the Darwinian Dilemma does not go through.²⁴

There is a second and more direct way of showing that OA is compatible with TT. Suppose that OA is true and also that an explanation in terms of OA is incompatible with TA. That is, suppose that if one can explain our tendency to hold

²⁴ Interestingly enough, Sinclair (2012) reaches a similar conclusion, namely that causing certain behaviors and representing certain states of affairs are compatible explanations. Nonetheless, he surprisingly denies that these ideas can be used to defend the Tracking Account.

Likewise, at some point Copp (2008) seems to suggest that the Output Account and the Tracking Account might not be incompatible, but his approach significantly differs from the one suggested here. First, his main purpose is not to defend the Tracking Account but what he calls a 'quasi-tracking thesis', which is much weaker (Street 2008, p. 211). Accordingly, he does not directly address Street's arguments against tracking accounts. Secondly, while he uses his society-centered theory in order to defend the quasi-tracking account, my proposal is much more general and is compatible with any realist approach. Finally, he pursues a very different strategy: whereas my defense appeals to evolutionary theories of meaning, he intends to show that the kind of facts that would enhance reproductive fitness (very roughly, satisfying the needs of the society) are approximately the same facts that would be tracked according to the society-centered account (Copp 2009). So Copp's reply merely insists on the idea that the content of our moral beliefs coincides to a large extent with the effects that are fitness-enhancing. Unfortunately, this reply might be insufficient for blocking Street's Dilemma. After all, Street (2006) agrees on this coincidence; her challenge is to provide an *explanation* of that fact (Enoch 2010) and Copp does not seem to offer one (Street 2008, p. 214).

certain moral beliefs by appealing to the behavior they elicited, an explanation in terms of its truth is excluded. Street assumes that a rejection of TA implies the falsity of TT. In other words, she takes for granted that only if the mechanism underlying moral beliefs were selected *because of its truth* could the tracking thesis be vindicated. However, this assumption can also be questioned.

Consider the teleosemantic model put forward by Dretske (1988, p. 84; see also Vicente 2012), which slightly differs from the one presented earlier. To see how it works, suppose a neuronal state R correlates with a state S. According to most philosophers, a mere correlation is insufficient for representation, so at this point R does not represent S. However, imagine that by chance R is linked to a motor response M, which turns out to be adaptive when S is the case. As a result, R is recruited by natural selection²⁵ and acquires the function of indicating S and the function of causing M. On Dretske's model, when R acquires the function of indicating S, it becomes a representation. What is interesting about this framework is that the adaptiveness of the motor response gives rise to two functions at the same time: R acquires the function of indicating S (and, in this way, becomes a representation) and it also acquires the function of causing M. So we first have a state R that causes an adaptive behavior M, and only when this mechanism has been selected for, R becomes a representation. Crucially, note that in that model the *truth* of R does not contribute to an explanation of the system (indeed, R was not even a representation before this structure was selected for). So this is a case in which the truth of the representation does not play any role in explaining the selection of the mechanism, and nonetheless R comes to track S. It is precisely the fact that R led to fitness-enhancing behavior what explains that R became a representation. Thus, this approach suggests that both OA and TT are fully compatible.

Let us move to the second reason why teleosemantics shows that premise 6 has not been established. I previously agreed with Street that TA alone cannot explain why a representational mechanism evolved. For instance, the fact that activation in the cricket's efferent neurons correlates with a predator being around does not explain why the mechanism evolved, unless we also consider the behavior of the receiver system. In general, one should not expect that merely representing truly provides any advantage to organisms (would perceptual experiences have benefited our ancestors if we did not have motor systems?). Representing certain features is only fitness-enhancing because they are coupled to mechanisms that lead us to act (or dispose us to act) in certain ways. Therefore, when Street rejects TA because it is not obvious why representing truly can be fitness-enhancing, she is putting forward a requirement that no representational system would satisfy. TA alone never provides a satisfactory explanation of the evolution of any representational system. Therefore, the fact that TA cannot account for the evolution of moral cognition on its own does not suffice for rejecting it, since a parallel reasoning would lead us to rule out a tracking account for any representational system.

²⁵ According to Dretske, the selection process giving rise to functions need not be natural selection. Learning, for instance, could be considered a selection process generating functions (see also Artiga 2010; Kingsbury 2008; Millikan 1984).

Finally, it should be mentioned that Street (2006, pp. 132–133) raises three additional challenges against TA: (1) ‘how does tracking account explain the remarkable coincidence that so many of the truths it posits turn out to be exactly the same judgments that forge adaptive links between circumstance and response (...)’, (2) how can it explain the fact that there are so many moral judgments that we could hold but do not and (3) what does the tracking account have to say about our observed predispositions to make false judgments (e.g. that we should favor in-group members)? Street assumes OA can satisfactorily explain all these facts, whereas TA can not. However, since I just argued the supporter of TA can and should accept both TA and OA, the very same answers she provides are available to the tracking theorist. Thus, I doubt these questions raise any additional worry for TA.²⁶

It is worth stressing that my arguments against premise 6 of DD do not depend upon the claim that the teleosemantics’ account is wholly correct.²⁷ The Darwinian Dilemma requires the falsity of TT and Street’s only evidence against it is the plausibility of OA. Teleosemantics is one way of showing that the truth of OA is compatible with TT, but there might be other ways of developing the same idea.

In conclusion, I think a careful look at evolutionary theories of representation show the abductive argument to be faulty. Crucially, and in contrast to previous metaethical discussions that have appealed to naturalistic theories of content, I am not using these theories in order to show that MORAL REALISM is true (see Harms 2000; cfr. Joyce 2001b), but merely to argue that a powerful argument against MORAL REALISM is flawed.

3 Conclusion

Summing up, the arguments purporting to show that the OA is much more plausible than the Tracking Account are unconvincing. None of the three arguments I considered is successful, so the truth of the Output Account does nothing to show that the Tracking Thesis is false or implausible. Furthermore, my rejection of the arguments did not rely on any particular positive approach of what moral properties are. Consequently, most moral realists can use the arguments defended here in order to defend the viability of the Tracking Thesis.

How does that bear on the Darwinian Dilemma? At least, it strongly suggests that premise 6 is far from established. Hence, unless further evidence for this claim is provided, the Darwinian Dilemma does not go through. The mere fact that evolutionary forces had a tremendous effect in shaping our tendency to hold certain

²⁶ In relation to the third question, one could object that teleosemantics cannot account for the existence of representational mechanisms that systematically produce false representations. An answer to this worry can be found in Artiga (2013).

²⁷ In that respect, it parallels Dowell’s (forthcoming) use of naturalistic theories of content. She employs teleosemantics in order to defend moral realism from the Moral Twin Earth Argument (Horgan and Timmons 1992a, b) and I use it to defend it from the Darwinian Dilemma. Similarly, our arguments do not require teleosemantics to be wholly correct.

moral beliefs does not show that they are unjustified. Our inclinations towards certain moral beliefs were probably favored by natural selection because they led our ancestors to behave in fitness-enhancing ways, but this claim should not diminish our confidence in them. If the arguments of this paper are on the right track, (A), (B) and (C) are fully compatible. The Darwinian Dilemma can not show that we should distrust MORAL REALISM. On the contrary, I hope I have provided some convincing reasons for thinking that we should distrust the Darwinian Dilemma.

Acknowledgments I would like to thank Eduardo García-Rodríguez, Maximiliano Martínez, Miguel Ángel Sebastian, Ljubomir Stevanovic, Moisés Vaca, Rodrigo Valencia, an anonymous referee and the audiences of the Seminario de Biología (PhiBio) and the Seminario Interuniversitario de Metodología Filosófica (SIMeF) for helpful criticisms and suggestions. This work was possible thanks to the Programa de Becas Postdoctorales de la UNAM.

References

- Artiga, M. (2010). Learning and selection processes. *Theoria*, 25(2), 197–210.
- Artiga, M. (2013). Reliable misrepresentation and teleosemantics. *Disputatio*, 37(5), 265–281.
- Artiga, M. (2014). Signaling without cooperation. *Biology and Philosophy*, 29(3), 357–378.
- Blackburn, S. (1993). *Essays in quasi-realism*. Oxford: Oxford University Press.
- Brink, D. (1989). *Moral realism and the foundations of ethics*. Cambridge: Cambridge University Press.
- Cao, R. (2012). A teleosemantic approach to information in the brain. *Biology and Philosophy*, 27(1), 49–71.
- Clarke-Doane, J. (2012). Morality and mathematics: The evolutionary challenge. *Ethics*, 122(2), 313–340.
- Comer, C., & Leung, V. (2004). The vigilance of the hunted: Mechanosensory visual integration in insect prey. In F. R. Prete (Ed.), *Complex worlds from simpler nervous systems* (pp. 313–335). Cambridge: MIT Press.
- Copp, D. (2008). Darwinian skepticism about moral realism. *Philosophical Issues*, 18(1), 186–206.
- Copp, D. (2009). Toward a pluralist and teleological theory of normativity. *Philosophical Issues*, 19(1), 21–37.
- Dowell, J. (Forthcoming). The metaethical insignificance of moral twin earth. In Russ Shafer-Landau (Ed.), *Oxford Studies in Metaethics*, Vol. 11. Oxford: Oxford University Press.
- Dretske, F. (1988). *Explaining behavior*. Cambridge: MIT Press.
- Dretske, F. (1995). *Naturalizing the mind*. Cambridge: MIT Press.
- Enoch, D. (2010). The epistemological challenge to metanormative realism: How best to understand it, and how to cope with it. *Philosophical Studies*, 148(3), 413–438.
- FitzPatrick, W. (2015). Debunking evolutionary debunking of ethical realism. *Philosophical Studies*, 172(4), 883–904.
- Fraser, B. (2010). Adaptation, exaptation, by-products and spandrels in evolutionary explanations of morality. *Biological Theory*, 5(3), 223–227.
- Fraser, B. (2014). Evolutionary debunking arguments and the reliability of moral cognition. *Philosophical Studies*, 168(2), 457–473.
- Gibbard, A. (1990). *Wise choices, apt feelings: A theory of normative judgment*. Cambridge: Harvard University Press.
- Godfrey-Smith, P. (1996). *Complexity and the function of mind in nature*. Cambridge: Cambridge University Press.
- Harman, G. (1977). *The nature of morality: An introduction to ethics*. Oxford: Oxford University Press.
- Harms, W. (2000). Adaptation and moral realism. *Biology and Philosophy*, 15(5), 699–712.
- Hauser, M. (2006). *Moral minds: How nature designed our universal sense of right and wrong*. New York: Harper Collins.
- Horgan, T., & Timmons, M. (1992a). Troubles on moral twin earth: Moral queerness revived. *Synthese*, 92, 221–260.

- Horgan, T., & Timmons, M. (1992b). Troubles for new wave moral semantics: The 'Open Question Argument' Revived, *Philosophical Papers XXI*, 153–175.
- Joyce, R. (2001a). *The myth of morality*. Cambridge: Cambridge University Press.
- Joyce, R. (2001b). Moral realism and teleosemantics. *Biology and Philosophy*, 16(5), 723–731.
- Joyce, R. (2006). *The evolution of morality*. Oxford: Oxford University Press.
- Joyce, R. (2013). The many moral nativisms. In Sterelny et al. (Eds.), *Cooperation and its Evolution*, Cambridge: MIT Press.
- Kahane, G. (2011). Evolutionary debunking arguments. *Noûs*, 45(1), 103–125.
- Kim, J. (1998) *Mind in a Physical World*, MIT Press.
- Kingsbury, J. (2008). Learning and selection. *Biology and Philosophy*, 23, 493–507.
- Kitcher, Ph. (2005). Biology and ethics. In David Copp (Ed.), *The oxford handbook of ethical theory*. Oxford: Oxford University Press.
- Mackie, J. L. (1977). *Ethics: Inventing right and wrong*. Harmondsworth: Penguin.
- Martinez, M. (2013). Teleosemantics and indeterminacy. *Dialectica*, 67(4), 427–453.
- Mason, K. (2010). Debunking arguments and the genealogy of religion and morality. *Philosophy Compass*, 5(9), 770–778.
- Millikan, R. (1984). *Language, thought and other biological categories*. Oxford: Oxford University Press.
- Neander, K. (1991). Functions as selected effects: The conceptual analyst's defense. *Philosophy of Science*, 58(2), 168–184.
- Neander, K. (1995). Misrepresenting and Malfunctioning. *Philosophical Studies*, 79(2), 109–141.
- Neander, K. (2012) Teleological Theories of Mental Content, *Stanford Encyclopedia of Philosophy*.
- Neander, K. (2013). Toward an informational teleosemantics. In J. Kingsbury, D. Ryder, & K. Williford (Eds.), *Millikan and her critics*. Oxford: Blackwell.
- Papineau, D. (1987). *Reality and representation*. Oxford: Blackwell.
- Price, C. (2001). *Functions in mind: A theory of intentional content*. Oxford: Oxford University Press.
- Ruse, M. (1996). *Evolution and religion: A dialogue*. New York: Rowman & Littlefield.
- Ruse, M. (2005). The darwinian revolution, as seen in 1979 and as seen twenty-five years later in 2004. *Journal of the History of Biology*, 38(1), 3–17.
- Schafer, K. (2010). Evolution and normative scepticism. *Australasian Journal of Philosophy*, 88(3), 471–488.
- Schulte, P. (2015). Perceptual representations: A teleosemantic answer to the breadth-of-application problem. *Biology and Philosophy*, 30(1), 119–136.
- Shafer-Landau, R. (2003). *Moral realism: A defence*. Oxford: Oxford University Press.
- Shafer-Landau, R. (2012). Evolutionary debunking, moral realism and moral knowledge. *Journal of Ethics and Social Philosophy*, 7(1), 1–37.
- Shea, N. (2007). Consumers need information: Supplementing teleosemantics with an input condition. *Philosophy and Phenomenological Research*, 75(2), 404–435.
- Sinclair, N. (2012). Metaethics, teleosemantics and the function of moral judgements. *Biology and Philosophy*, 27(5), 639–662.
- Stegmann, U. (2009). a consumer-based teleosemantics for animal signals. *Philosophy of Science*, 76(5), 864–875.
- Sterelny, K., Joyce, R., Calcott, B., & Fraser, B. (Eds.). (2013). *Cooperation and its evolution*. Cambridge: MIT Press.
- Street, Sh. (2006). A Darwinian Dilemma for realist theories of value. *Philosophical Studies*, 127(1), 109–166.
- Street, Sh. (2008). Reply to Copp: Naturalism, normativity, and the varieties of realism worth worrying about. *Philosophical Issues*, 18(1), 207–228.
- Vavova, K. (2015). Evolutionary debunking of moral realism. *Philosophy Compass*, 10(2), 104–116.
- Vicente, A. (2012). Burge on representation and biological function. *Thought*, 1(2), 125–133.
- Wilkins, J., & Griffiths, P. (2013). Evolutionary debunking arguments in three domains: Fact, value, and religion. In James Maclaurin & Greg Dawes (Eds.), *A new science of religion*. London: Routledge.