

Moral Realism and Evolution

by

Matthew Scarfone

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Arts

in

Philosophy

Carleton University
Ottawa, Ontario

© 2013, Matthew Scarfone



Library and Archives
Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

ISBN: 978-0-494-94598-8

Our file Notre référence

ISBN: 978-0-494-94598-8

NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protègent cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

The evolutionary challenge to moral realism is the challenge of explaining the possibility that our moral intuitions are true given that those intuitions are evolutionarily advantageous. If natural selection tracks fitness-enhancement, and not necessarily truth, then we should not be confident that our moral beliefs are true. Or so the evolutionary challenge suggests.

Chapter 1 gives the distinction between Moral Realism and Anti-realism, and suggests that we should be Anti-realists. Chapter 2 is an evolutionary account of our ability to make moral judgments, and an exposition of the evolutionary challenge. Chapter 3 details Moral Realists who have argued that, even if our moral beliefs are evolutionarily advantageous, we have an additional reason to suppose that our moral beliefs are true. I argue that these *additional reasons* are insufficient. And Chapter 4 outlines two idiosyncratic accounts of Moral Realism that attempt to salvage the metaethical position. Here too the arguments fail, and I show how any further attempts at preserving Moral Realism in light of the evolutionary challenge are similarly doomed.

Table of Contents

Abstract	ii
Table of Contents	iii
Chapter 1: Moral Realism and Anti-realism	5
1.1 Geocentrism/Heliocentrism and Moral Realism/Anti-realism	5
1.2 Why We Should be Anti-realists	6
1.3 Metaphysical Naturalism	9
1.4 Setting up the next chapters	14
Chapter 2: The Evolutionary Challenge to Moral Realism	15
2.1 Group-living and Our Moral Sense.....	15
2.2 Simple Rules and Signaling	19
2.3 Reciprocal Altruism	20
2.4 Other Ways of Helping	23
2.5 Categorical Requirements.....	25
2.6 Moral Properties.....	28
2.7 The Darwinian Dilemma.....	30
2.8 Raising and Addressing Berker's Worry.....	34
2.9 The Basic Moral Realist Response	37
2.10 Summary.....	38
Chapter 3: Why Should We Be Moral Realists?	40
3.1 A Quick Look at Moral Realism.....	40
3.2 An Additional Reason for Moral Realism	41
3.3 The Platonists.....	42

3.4	Mathematics and Morality	45
3.5	Social and Cultural Influence.....	52
3.6	The Theist	56
3.7	Medium-sized Dry Goods in our Immediate Vicinity	59
3.8	Changing Moral Beliefs.....	61
3.9	The Similarity between Moral Realism and Theism	65
3.10	Summary.....	68
Chapter 4: Critiquing Moral Realist Responses		70
4.1	Morality is Objective	70
4.2	Is Morality Mind-independent?	75
4.3	Preserving Mind-independence	79
4.4	Basic and Advanced Moral Beliefs.....	83
4.5	Problems with the Moral Realist Response	86
4.6	False Hope for Moral Realism.....	87
4.7	Our Basic Beliefs are probably false	89
4.8	Our Advanced Beliefs are built from Basic Beliefs.....	91
4.9	Why be a Moral Realist?	93
References		94

Chapter 1: Moral Realism and Anti-realism

1.1 Geocentrism/Heliocentrism and Moral Realism/Anti-realism

We must no longer say upon these principles that fire heats, or water cools, but that spirit heats, and so forth. Would not a man be deservedly laughed at, who should talk after this manner? I answer, he would so; in such things we ought to think with the learned, and speak with the vulgar. They who to demonstration are convinced of the truth of the Copernican system do nevertheless say the sun rises, the sun sets, or comes to the meridian; and if they affected a contrary style in common talk, it would appear very ridiculous.

– George Berkeley (1710 [2009], 45)

Buckminster Fuller, like Berkeley, noted a few anachronistic features of the modern English language. Today we still talk of *sunrises* and *sunsets*. Of course, the Sun does not actually rise or set, rather the Earth revolves around the Sun, giving us the impression that the Sun is rising and setting. While everyone can now recognize this fact, we have nevertheless retained the terms *sunrise* and *sunset* from our pre-Copernican geocentric-biased past. Fuller's proposed replacement neologisms are the words *sunlight* and *sunclipse*. Likewise, he suggested dropping most usages of the terms *up* and *down*. He felt that these words make sense only if we have a planar concept of direction, but are inconsistent with the fact that we exist on a planet in Space, where *up* and *down* are essentially meaningless. Fuller's alternatives here are *in* and *out*, where the reference point is the gravitational centre of Earth. As opposed to going *upstairs*, one should say they are going *outstairs* - that is, moving further out from Earth's centre. And to round things out, Fuller felt that *world-around* should replace *world-wide*, since the latter is leftover from flat-Earth thinking that died out centuries ago.

I mention Fuller's suggestions here not in the hope of pursuing some alteration to particular linguistic conventions, rather I am intrigued by the way we can reflexively

think one thing yet *know* another. As Berkeley says, we should think with the learned, yet speak with the vulgar. What follows here, however, is my attempt to speak with the learned.

I think there are interesting parallels between the above anachronisms and how we today think and talk about morality. In particular, it seems that we often both think and talk of actions, events, and behaviours being good or bad, moral or immoral, just or unjust, and so on. But can something *actually* be good or bad, moral or immoral, just or unjust? And why do we reflexively think that things can? Though the Sun does not actually rise or set, it is obvious why we are tempted to think so: from our vantage point on Earth, the Sun appears to do these very things, and so pre-Copernican humanity is easily forgiven for thinking that the Sun *actually* rose and set. As said, my focus here is on how we often think and speak of actions or events being good or bad, or moral or immoral, and just as the Sun does not *actually* rise or set, so too does it seem to me that nothing is *actually* moral or immoral. Roughly speaking, and somewhat imprecisely, my position is moral anti-realist and its opposite moral realist.

1.2 Why We Should Be Anti-realists

Throughout the thesis I focus exclusively on so-called *non-naturalist moral realism*. This version of realism makes the following claims: first, that in making moral statements we reflexively believe that the statements are truth-apt (i.e., capable of being true or false); second, that some of our moral statements are in fact true; and third, that what makes our moral statements true is something non-natural. I'll add to this brief description of non-

naturalism in (2.7) and (3.1), but the reader should take all references to “moral realism” as meaning “non-naturalist moral realism”.

In the following chapters I argue that it is tempting to think like moral realists, but that this inclination is as misguided as thinking that the Sun actually rises and sets. And just as we can offer a convincing account of why we would be tempted to think and speak of the Sun rising and setting, there is also a convincing account for why we are inclined to think and speak of actions, events, or behaviours as being actually moral or immoral. Specifically, I link this latter inclination towards moral realism to developments in our evolutionary history. It is plausible that we have developed cognitive mechanisms that result in our having a disposition to view reality through the lens of moral realism. I suggest that if these cognitive mechanisms are the result of evolutionary processes, then it is an open question whether they evolved to track evolutionary fitness, truth, or both. I argue that the mechanisms evolved with respect to evolutionary fitness, and thus we may undermine the justification for moral realism. That is, if we can give a convincing account of why we are inclined to think in realist terms, an account that makes no reference to *actual* moral truths, then we can no longer use the fact that we reflexively think and speak like moral realists as a justification of that metaethical position.

What my argument will do is make the moral realist position much more difficult to present. If there is indeed a dearth of facts that support the moral realist framework, and if there is a credible explanation to be given for why we would come to have realist dispositions, then the onus is shifted squarely onto the realist to explain either why we should nevertheless suppose a realist framework, or to explain why the explanation I give for our realist inclinations is unsatisfactory. As will become apparent, much of this

discussion will turn on what exactly is going to count as a relevant fact. Some moral anti-realists think that the only facts that exist are natural facts (i.e., metaphysical naturalism), whereas some moral realists argue for the existence of non-spatiotemporal facts in addition to natural facts (i.e., metaphysical platonism¹, non-naturalism, or supernaturalism). I argue that the difference between these two positions, between metaphysical naturalism and metaphysical platonism, is ultimately insurmountable.

My argument here is an inference to the best explanation, and it proves to be a powerful tool against the possibility of moral realism. In his (2007) Terence Cuneo argues for a version of moral realism that supposes that non-natural facts exist (i.e., non-naturalist moral realism). In doing so Cuneo presents a largely defensive account, mostly attempting to argue *against* anti-realism rather than *for* realism. The reason he gives for this strategy is that, given our exceedingly strong intuitions regarding moral realism, “it is extraordinarily challenging to find other ways to argue for [realist] views in a non-question-begging fashion” (2007, 11). Cuneo is here suggesting that moral realism gains immense plausibility as a metaethical position because it seems as natural and obvious to us as the existence of the external world, or (as I said earlier) the rising and setting of the sun. Cuneo uses our intuitive experiences of morality (i.e., our moral phenomenology) as *prima facie* justification for the truth of moral realism. And because he leans heavily on our moral phenomenology, Cuneo’s positive argument for moral realism is likely to be fairly thin, and thus blocking sceptical arguments for anti-realism (like blocking external world scepticism) is the easier of the two tasks.

¹ Following the literature, I use a lowercase-p when discussing metaphysical (or mathematical) platonism in order to distinguish from capital-p Platonism (i.e., views attributable directly to Plato).

I agree with Cuneo that moral realism is an immensely intuitive meta-ethical position. However, what Cuneo overlooks is *why* we have these seemingly obvious views when it comes to morality. I argue that if we have a plausible account of why we have our moral intuitions, in a way that makes no reference to the truth of those intuitions, then we have a *prima facie* reason for doubting our moral intuitions. If this conditional holds, then we should not be able to rely on the unreflective plausibility of our moral intuitions in support of the moral realist position. In effect, our moral phenomenology would suddenly lose its justificatory force. And if this force is lost, then it seems that the moral realist will indeed be required to present a positive account of moral realism. If this proves to be too difficult, as Cuneo suspects, then so much the worse for moral realism.

1.2 Metaphysical Naturalism

As said, metaphysical naturalism is the position that the only things that exist are natural things, things that exist spatiotemporally, things we can understand scientifically (via *methodological naturalism*), and so on. *Natural things* is here contrasted with the existence of non-spatiotemporal things, i.e., non-natural or supernatural things. This latter position is what I will refer to as *metaphysical platonism*. Platonism, in this sense, is the position that there truly exist non-spatiotemporal objects. So we can understand naturalism as the denial of all forms of platonism. According to metaphysical naturalists, if an object exists, it exists in a spatiotemporal sense.

In his defense of naturalism, Richard Carrier (2005) writes that

all metaphysical naturalists believe that if anything exists in our universe, it is a part of nature, and has a natural cause or origin, and there is no need of any other explanation (67).

Carrier's characterization of naturalism is that it is an inference to the best explanation. When we suppose that the universe is natural, we need not posit superfluities, nor give problematic accounts of how spatiotemporal human beings can relate to non-spatiotemporal objects. On the other hand, if the universe *did* contain non-spatiotemporal objects then there is a deep worry about how we can have knowledge of them, or even interact with them in any way.² I take it that our knowledge of natural things is relatively uncontroversial (or at least *less* controversial), whereas knowledge of non-spatiotemporal things remains problematic. Certainly this is not a knockdown argument against metaphysical platonism, but it shows one well-known and often discussed problem with the position. Naturalism can easily skirt this problem by striking non-spatiotemporal objects from our ontology. There is no *prima facie* problem of how we can gain knowledge of natural facts, while there is a *prima facie* problem of how we can gain knowledge of non-natural facts.

Though he doesn't use the terms "metaphysical naturalism", "methodological naturalism", or "metaphysical platonism", Allan Gibbard (2012) presents a quick and cogent defense of naturalism. He suggests that "with each passing decade we get further indications that naturalistically is the way we must learn to understand ourselves" (24). By "naturalistically", I take it that Gibbard means *in ways presumed by metaphysical naturalism*, and by "further indications" I understand him to have in mind the success of

² This so-called *Benacerraf Problem* in philosophy of mathematics, a problem that perhaps generalizes to all realist frameworks. Paul Benacerraf (1975) writes that "something must be said to bridge the chasm, created by... [a] realistic... interpretation of mathematical propositions, between the entities that form the subject matter of mathematics and the human knower" (675). Plainly, if mathematical propositions refer to non-spatiotemporal objects, how can spatiotemporal human beings know about these objects? I have more to say about this problem in (3.4).

methodological naturalism (i.e., our best scientific accounts). Gibbard thinks that arguments or suggestions that lead us to positions “that we don't know how to reconcile with the view that human phenomena are natural” prove problematic (24). These irreconcilable or incompatible positions are those posited by platonists, who suppose that something non-natural or non-spatiotemporal exists. However, to date we have no confirmation of any non-natural or supernatural objects, so we have yet to confirm that metaphysical platonism is a plausible metaphysical position. As such, it seems that currently we are only justified in presuming that the universe is naturalistic. I hold the acceptance of metaphysical naturalism (and the denial of metaphysical platonism) as one of my background commitments here.

So given the successes that issue from methodological naturalism, why would anyone be motivated to deny metaphysical naturalism? If I'm to speculate, it seems that some feel that the best way to accommodate or make sense of certain domains is by positing non-natural or supernatural entities as part of the universe. For example, in order to explain how we can make true (or seemingly true) mathematical claims, one might suppose that numbers and mathematical operators are real objects. When we say that $2+2=4$, one might suppose that “2”, “4”, “+”, and “=” are mathematical objects that *actually* exist, and that our mathematical sentences are simply true claims about these mathematical objects. But clearly numbers are obviously not real in the *reach out and touch them* sense. So one might be tempted to say that numbers are real not in a natural sense, but that they are real in a non-natural sense. This is *mathematical platonism*.

Another domain that may motivate a denial of metaphysical naturalism comes from trying to account for moral statements. One might suppose that some moral

statements, like *murder is morally bad*, are truth-apt (i.e., capable of being true or false). But what exactly can make our moral statements truth-apt? Here one might suppose that so-called *moral facts* somehow confer truth onto our moral statements: that there are moral properties that our moral statements *actually* possess. Here the parallel with mathematics is helpful: just as our mathematical statements are true when they correctly describe mathematical objects, so too might our moral statements be true when they correctly describe moral facts. But again, some suggest that moral facts are not real in the *reach out and touch them* sense, but that such facts are real in the same way that mathematical platonists suppose that numbers are real. For some moral realists, moral facts are thought to be non-spatiotemporal or non-natural. So, certain issues, say in mathematics and morality, are thought by some to be best solved by supposing that non-spatiotemporal objects exist. And the existence of such objects leads to a denial of metaphysical naturalism.³

I want to look briefly at the platonist strategy. As I see it, the plausibility for platonism stems from an attempt to make sense of our intuitions. When we utter moral sentences we certainly do seem to predicate certain properties of subjects – we have the experience that murder *really is* morally bad, or that *moral badness* is something that *murder* really possesses. We should say that morality presupposes cognitivism, which is the view that our beliefs are truth-apt. In turn, cognitivism presupposes some form of moral predication. So we should say that morality ordinarily presupposes moral predication. But how can we make sense of such predication? One way, as said, is to

³ I have more to say about these forms of platonism in (3.4).

posit that there are non-spatiotemporal objects that (somehow) render our moral statements true or false.

However, note that this strategy supposes that our metaphysics must accommodate our intuitions about what we take ourselves to be doing in certain domains. If, for example, we think we predicate moral properties of actions when we utter moral statements, the metaphysical platonists say that our ontology must really contain those properties. And, platonists continue, if the properties must be non-spatiotemporal or non-natural, then our ontology must contain non-spatiotemporal or non-natural things. So metaphysical platonists think that our beliefs or intuitions about what we are doing when we utter moral statements takes priority, and that our ontology should conform to our intuitions about what we are doing when we make moral claims.

This way of doing metaphysics seems to me entirely backwards. Rather than fitting our metaphysics to our pre-theoretic intuitions, I think we can figure out why we have such intuitions in a way that fits our metaphysics. That is, in contrast to platonists, I do not suppose that our intuitions take priority. Instead, I think that we can make sense of our intuitions by looking at our evolutionary history, and through this naturalistic understanding of how we came to have our intuitions in certain domains we can undermine the motivation towards platonism. If I am right, then we do not need to posit the existence of non-spatiotemporal or non-natural objects in order to make our ontology fit our intuitions. Rather we should do just the opposite: we should instead understand our intuitions in a way that fits with our ontology. Doing so, I argue, makes certain metaethical positions unpalatable; the first target, of course, is non-naturalist moral realism.

Obviously I am assuming here that our ontology is best described by our current science. Regarding this, I am in agreement with Gibbard that all our available evidence reveals that the best way to understand ourselves, and our place within the universe, is naturalistically. Should future research show the plausibility of non-spatiotemporal objects or non-natural facts, at the time we would be justified in presuming metaphysical platonism. So presently the discussion between metaphysical platonists and metaphysical naturalists forks: we either upend our entire ontology in order to accommodate our intuitions on topics like mathematics, morality, etc., or we come to understand our intuitions in those domains naturalistically in a way that preserves the naturalist ontology revealed by methodological naturalism. I defend the latter option.

1.4 Setting up the next chapters

In the next chapter I introduce the core issues necessary to understand my thesis. First, I give an account of why we are inclined to think like moral realists that turns on our evolutionary history. Second, I explicate two versions of the evolutionary challenge to moral realism, and in doing so give a brief summary of moral realism as a metaethical position. And third, I outline the basic way that moral realists have attempted to answer the evolutionary challenge.

Chapter 2: The Evolutionary Challenge to Moral Realism

In this chapter I give a naturalistic account of our moral intuitions. I argue that some of our moral intuitions, especially those that strike us as immediately plausible, were selected for (or not selected against) because of the fitness-enhancing benefits conferred on those that had them. Being able to think that some things are *actually* good or bad is what I refer to as *having a moral sense*. What does a naturalistic interpretation of the origin of our moral sense say about plausibility of moral realism as a metaethical position? I say that if our moral sense is best understood as responding to fitness-enhancement, then we are not justified in thinking that our moral intuitions necessarily track truth. This is the so-called *evolutionary challenge to moral realism*. I present three versions of the challenge: first I give an early and accessible version from Richard Joyce, and highlight some problems therein; next I present Sharon Street's *Darwinian Dilemma*, which seems to me the strongest account in the literature; and finally I offer my own, which is a qualification of Street's given some recent criticisms raised by Selim Berker. I end the chapter by briefly describing the way that moral realists have attempted to respond to the evolutionary challenge.

2.1 Group-living and Our Moral Sense

At the heart of the evolutionary challenge to moral realism is the suggestion that natural selection has “provide[d] us with the tendency to invest the world with values” (Joyce 2001, 135). The suggestion here is that human beings evolved in such a way that we are inclined to characterize the world using moral concepts – i.e., *this action* is right or

wrong, or *that event* is good or bad, etc. At some point in our evolutionary development, we evolved to come to see actions or events as containing moral properties. By contrast, most (very likely *all*) non-human animals do not characterize the world in this way. We can shorten this by saying that human beings have a *moral sense*, which means we are able to characterize actions, behaviours, etc. as containing or lacking moral properties, or as being moral or immoral, good or bad, just or unjust, and so on. So the evolutionary challenge begins by probing why or how human beings could have come to do this. Why do we have a moral sense in the first place? I offer a naturalistic account here.

Brute existence is difficult stuff, surviving *and* reproducing more difficult still. The deck of nature is stacked against particular individuals surviving long enough to pass on their genes. One needs to find food, and ward off enemies, both visible (predators, etc.) and invisible (diseases, famines, etc.). One usually needs to find a mate, couple, and hope or help the offspring reach sexual maturity themselves. There is a lot of luck involved in all of this, though there are certain strategies, for lack of a better word, that can tilt the balance towards survival and reproduction, if only slightly.

One such strategy is communal or group living, which is the norm for primates.⁴ Why live in groups? Because cooperation proves advantageous. Hans Kummer (1978) suggests that “the essential benefit of social life is cooperation, its essential disadvantage competition” (43). Similarly, John Rawls (1971) writes:

[A]lthough a society is a co-operative venture for mutual advantage, it is typically marked by a conflict of interest as well as by an identity of interests. There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to live solely by his [or her] own efforts. There is a conflict of interests since persons are not indifferent as to how the greater

⁴ Group living also confers benefits for some other animals. See Marc Bekoff and Jessica Pierce’s (2011) Mark Bekoff’s (2007).

benefits of their collaboration are distributed, for in order to pursue their ends they each prefer a larger to a lesser share (4).

Living in groups gives us potential partners and allies for, say, finding food, building shelters, warding off enemies, finding mates, and so on – things we very much require for our own survival. But group living also puts us in close proximity to other individuals who need those same things for their own survival. As Rawls notes, others are not indifferent to how collaboration unfolds, and all else being equal we would each prefer a larger share to a lesser one. There is a profound tension, then, between the benefits conferred by group living and the drawbacks.

How can we ease that tension? One way is through morality. Among other things, we use morality to get along with others: we use it to know which persons to praise, and which to punish. And we use morality as justification for our decisions – we might say to someone, “*but you can't φ! φ-ing is immoral!*” But on the face of it, morality is an anti-evolutionary force. If biological existence is thought to be governed by struggle, conflict, self-interest, and so on, all with the goal of passing on our genes, then morality does not seem to be readily captured therein. As we ordinarily understand it, morality is a way of *reducing* struggle, *easing* conflicts, and *eliminating* self-interest. Likewise, morality does not seem to have much to do with genetic reproduction. Given this, how can we make sense of morality? Does it constitute a break from evolution, or can it be captured by evolutionary considerations?

I argue that morality is indeed a way to reduce struggle, ease tensions, and muffle (overt) self-interest, but by acting morally we thereby facilitate group living, and thus serve evolutionary ends. By such assuagement, individuals within groups are conferred with evolutionary advantages in comparison to those who (for whatever reasons) cannot

navigate group living. And it is because of the advantages that arise from acting morally that we can plausibly view the *prima facie* anti-evolutionary force of morality as having an evolutionary origin.

The question I address is what insights this account can offer with respect to a particular metaethical position. Given that morality originated and evolved naturalistically, what can we then say about the plausibility of non-naturalist moral realism, the position that moral claims are made true by something objective and mind-independent? In this chapter I outline a general account of the origins of our moral sense.

The story is roughly as follows:

- i) Human beings descended from non-human animals who lived in groups.
- ii) But there is a profound tension that comes from living in groups.
- iii) So there needs to be a way of managing this tension or else the disadvantages of group living may outweigh the advantages.
- iv) Morality is an effective means of minimizing that tension.

Note how this brief outline does not make any reference to the *truth* of our moral beliefs. As we will see, if we have no good reason to suppose that our moral beliefs are *true*, and a good reason to think we would have our moral beliefs *irrespective* of their truth, then moral realism is a problematic metaethical position because it needlessly posits that our moral beliefs are (and must be) true. If morality is simply a way of easing group tension so that the benefits of group living outweigh the disadvantages, then morality plausibly tracks evolutionary fitness (i.e., moral beliefs are fitness-tracking). And if we have no additional reason to suppose that morality also tracks truth, then we are not justified in thinking that our moral beliefs are actually true.

2.2 Simple Rules and Signaling

Frans de Waal posits that so-called *simple rules* emerge as a result of group-living. Simple rules govern how to behave by prescribing certain actions or behaviours and proscribing others (Flack and de Waal 2000, 3). These governing rules begin to emerge when human beings form groups together, as well as when non-human primates form groups. This is because of the benefits conferred on those that live in a group that operates according to rules – as said, we get valuable partners and allies (2.1).⁵ Of course, the individuals within the groups are concerned with survival. When one is better able to survive and reproduce, then one's *evolutionary fitness* can be said to increase, and when one is, for whatever reason, unable to do so, then one's evolutionary fitness decreases.

Simple rules, or social governing rules, are a way in which evolutionary fitness is served. These rules appear to generate and reinforce “connections among individuals... [and] facilitate co-operative social interaction” (Flack and de Waal 2000, 3). The way in which such connections are made is through the *commitments* that one makes to behave in a certain way. For primates, by accepting certain social rules, and more importantly by signaling to other primates that one has accepted them, an individual commits to “behave in ways that later may prove contrary to independent individual interests” (3). Such social rules appear to mediate the interactions of individuals by allowing them to lower their guard, to a certain extent. If primates signal to each other that they will behave in regular and non-threatening ways, then two things happens. First, individuals expend less energy and resources, spend less time intuiting if others have hostile intentions, expend less

⁵ We might say that a society or group with some rules is preferable to one that lacks them entirely, since this allows us to know that others will act in reliable ways.

energy protecting themselves (as well as kith and kin) from possible attacks, expend less energy securing goods from annexation. And second, individual primates gain immeasurably valuable partners and allies, partners and allies for finding food, getting mates, etc. de Waal (1996) writes:

The ability and tendency to construct such associations, and to seek security within them, are products of natural selection found in members of species with better survival chances in a group than in solitude. The advantages of group life can be manifold, the most important being increased chances to find food, defense against predators, and strength in numbers against competitors (9).

So commitment or signaling, and the social rules that enable it to happen, contribute to sociality; rules better equip us to live in societies. Social organization would be impossible without some kind of organizing principles – some rules are preferable to no rules.⁶ Social organization brings a host of benefits to the individuals who manage to get along with others. If social organization has this benefit, then the social rules that enable social organization can be seen as conferring benefits to the individual. Thus pro-sociality can be said to enhance one's evolutionary fitness.

2.3 Reciprocal Altruism

To better understand the advantages of pro-sociality we can look at the theory of *reciprocal altruism*. An altruistic behaviour can be defined as one that “benefits another organism... while being apparently detrimental to the organism performing the behavior” (Trivers 1971, 35). Robert Trivers offers the example of someone attempting to save a drowning person: this ostensibly benefits the drowning person (i.e., their survival chances

⁶ de Waal shows that primates will kiss and embrace after fights, and engage in other reconciliation behaviours (de Waal 2000). See also Flack and de Waal (2007) for other examples of signaling.

increase) while coming at a cost to the first person (i.e., they put themselves at risk of drowning).

How might an altruistic behaviour like this come about? Why would anyone bother to help the drowning person? Trivers considers three possibilities (36):

- i) there is a random dispensation of altruism
- ii) there is non-random dispensation by reference to kin
- iii) there is non-random dispensation by reference to the altruistic tendencies of the recipient

A bit of unpacking here: (i) suggests that we are altruistic towards others, without any further considerations – we randomly help various people, anywhere from helping everyone to helping no one: random dispensation means, of course, that altruism occurs irrespective of any particular end; (ii) suggests we are altruistic to those who are related to us, proportionally – we are more likely to help a sibling than a cousin, and a cousin than a stranger (this is kin-selection theory [more in 2.4]); and (iii) suggests that we are altruistic to those who we take to be likely to be altruistic themselves. (iii) is reciprocal altruism: we act altruistically based on the supposition that those we benefit with our altruistic actions are likely to be altruistic towards us, if such an opportunity should arise for them.

Recall the aforementioned ways in which cooperative behaviours prove beneficial: finding food, warding off enemies, finding a mate, and so on. Altruistic behaviours, provided by others, help us immensely in these situations. For example, if we need food, then food sharing by others is a desirable altruism-based option: the behaviour is altruistic since it comes at a cost to the giver (less food for them) while being beneficial to the receiver. Engaging in reciprocal altruism, then, contributes greatly to pro-sociality, and indeed is perhaps the basis of sociality. Given the argument presented in (2.1), we

can now say that being pro-social, in the form of reciprocal altruism, confers benefits to the individuals who engage in it. This constitutes an evolutionary pressure that favours those individuals who are prone to altruistic behaviours in the manner suggested by (iii), and thus an evolutionary pressure towards pro-sociality itself.

But why not avoid reciprocating? Why not accept the benefits of others while refraining from extending them any benefits in return? Trivers calls this *cheating*.⁷ Cheating is *prima facie* advantageous since one gains from the altruism of others, without the costs of having to act altruistically oneself. But while this is ostensibly beneficial, Trivers suggests that the cheater option will be selected against “if cheating has later adverse affects [sic] on [one’s] life which outweigh the benefit of not reciprocating” (36). It is easy to see how this could be so. Being identified as a cheater, as a person who may not or will not reciprocate helping behaviours, will likely entail that you are far less likely to receive the benefits of cooperation. If (iii) is correct, and altruistic behaviours are non-random by reference to the likelihood that the helped individual will be altruistic as well, then being identified as a cheater may entail that altruistic behaviours cease (or are much less likely) to come your way. Because of the immense benefits of reciprocal altruism, and of group-living, failing to reciprocate constitutes a maladaptive outcome, and thus stands as something that may be selected against. In this way, reciprocal altruism is selected for, and cheating is selected against.

⁷ He further identifies two types of cheating: *gross* and *subtle cheating*. When one grossly cheats, the cheater simply fails to reciprocate, and when one subtly cheats, the cheater reciprocates but gives less than expected (Trivers 1971, 46).

There is far more that could be (and has been) said about reciprocal altruism, and the conditions which make it likely to emerge and persist. For my purposes here, the cursory account presented should prove sufficient. Since I will be discussing the way in which the origins and evolution of morality make certain meta-ethical positions unjustified, what matters here is simply that morality originated and evolved, and not necessarily how exactly it did so.⁸

2.4 Other Ways of Helping

As seen, the origins of morality in human beings plausibly lie in the development of pro-sociality through reciprocal altruism. As Trivers seems to suggest, kin selection theory is insufficient for explaining the dispersal of altruistic actions. Kin selection, recall, suggests that altruism correlates with genetic relatedness. However, one problem is that the mechanisms posited (say, kin recognition) by which a person is reliably able to recognize that another person shares a certain percentage of genetic material with themselves, are unreliable. Imagine fraternal twins separated at birth: while they would of course share 50% of their genetic material, there would be no reliable means by which they could realize these shared genetics if they, by chance, crossed paths in the future. Because of these two things combined (one, the way in which kin selection under-represents the dispersal of helping behaviours; and two, the lack of a reliable mechanism

⁸ However, not all metaethical theories are readily captured by this kind of evolutionary account. For example, one might argue that the origins and evolution of morality are irreparably tied to which metaethical accounts are justified. Certain theistic accounts might suggest something along these lines. I offer (and subsequently critique) one such account in (3.6).

for discerning genetic relatedness), it seems to me that kin selection can be seen as a relatively incomplete explanation of the origins and persistence of cooperation.

Additional qualifications have proven helpful, namely cooperation by proximity. That is, we may be inclined to help those around us, those with whom we are raised and live, and those with whom we share important milestone markers (e.g., birthdays, rites of passages, etc.). This eliminates the problem of identifying genetic relatedness, since this is *no longer the relevant reference point* – if someone *seems* to be our kith or kin, this is enough to secure reliable helping behaviours. But this too under-represents helping behaviours, and moreover it is not the central issue that the evolutionary challenge addresses. The way in which helping behaviours are dispersed *beyond* kith and kin unto those who are not (and do not seem) related to us is the important question.

To this end, recognizing the relationship between simple rules and reciprocal altruism proves informative. Recall that simple rules quickly emerge within groups, and that signaling is one way in which we broadcast to others that we are committed to those rules. If (iii) is correct (i.e., if we are altruistic in reference to the altruistic tendencies of others) then we may be signaling our cooperative intentions to others in order to be able to benefit from their altruism. We do not attempt to blindly intuit whether or not someone else is likely to be altruistic themselves. Instead, we rely on signals, and this seems to be a more efficient way of satisfying the dispensation pattern suggested by (iii).⁹

⁹ There are legitimate questions about the exact nature of these signals. How expensive was their evolution? How are they maintained? And so on. As my focus is on the plausibility that our moral sense *did* evolve, and what that says for a particular meta-ethical view, I set aside these *how* questions.

2.5 Categorical Requirements

If the connection between signaling and reciprocal altruism has been established, we can now move on to the question of how altruistic behaviours are carried out. An important question here is: what is the mechanism or mechanisms that motivate us to help or cooperate with others? One answer is *desires*. Presumably, we have evolved to have certain desires to help others: as intensely social creatures, we are driven to assist others, and we want to be helped by others when we ourselves are in need. The above concerns about kin selection aside, human beings do display strong desires to help and be helped.

But how do our desires get the job done, as it were? How do desires motivate us to do the things that are evolutionarily advantageous? We might say that human beings are desire-maximizers: maybe we are constantly attempting to fulfill our desires, and constantly calculating the best way of doing so. Tied to this is the notion that our desires are perfectly correlated with what is good for us in an evolutionary sense. In following our desires, we thereby serve fitness-enhancing ends. But both suggestions here seem mistaken: human beings are *not* mere desire-maximizers, nor are our desires perfectly correlated with evolutionary fitness. Joyce (2001), for one, argues that desire-maximization in particular proves to be a fickle enterprise, as our desires are easily altered or overridden, and thus ultimately prove unreliable. Joyce thinks that if our helping behaviours were motivated *only* by attempts to maximize our desires, the mechanism would not reliably secure whatever behaviours need to be secured.¹⁰ Our desires are readily supplanted: for example, our desire to help a friend in need could temporarily be overridden by a desire to eat, or to sleep, or to have sex.

¹⁰ See Joyce's (2001: Chapter 6).

We can note that while such desire supplanting indeed happens all the time, desires themselves certainly are not necessarily fitness-tracking. This isn't to say that our desires are not intensely felt, and that they do not motivate us, but rather that behaviours motivated *solely* by desires are susceptible to being supplanted by *other* desires. And without some assurance that our desires are perfectly correlated with that which serves evolutionary fitness, it is plausible that our desires would in fact be supplanted by something else that could more reliably secure such ends.

Joyce posits that having our desires supplanted by a stronger sense of requirement leads to the behaviours in question being more likely to take place. He thinks that the concept of *moral requirement* is this stronger sense, and our believing that ϕ is *morally required* entails that (all things considered) we are more likely to ϕ . Joyce argues that

[a]n individual who thinks that it is morally repugnant to cheat [his or] her comrades is less likely to do so – perhaps even less likely than a person who sees that cheating will harm [his or] her own long-term interests – and thus on the assumption that cheating is frequently maladaptive, the moral belief may be selected for. The plausibility of this claim seems independent of whether cheating one's comrades (or anything else) is morally repugnant (2012, 3).

This argument, specifically the part assuming that cheating is often maladaptive, implicitly relies on the fact that reciprocal altruism is beneficial (2.3). Thinking that something is *morally required* makes it more likely that we will do it. Michael Ruse (1998) goes so far as to suggest that morality “simply does not work (from a biological perspective) unless we believe that it is objective” (162). Less strongly, Joyce seems to suggest that an objective sense of morality *better secures* certain behaviours. Joyce's suggestion, rather than Ruse's, seems to me more plausible. If we are compelled to help others because of a process less fickle than desire-maximization, then it is perhaps more likely that we will actually help others. The more reliable mechanism here is being able

to think in terms of *moral requirement*, or having a moral sense. And because helping behaviours are on average beneficial, having this moral sense contributes to reproductive fitness, and our fitness increases as a result of being able to think in moral terms.

So Joyce's claim is that if we think that we are morally required to help others, we are more likely to do so than if we did not feel morally required to do the same thing. But how could natural selection have cultivated in us a sense of moral requirement? This is a core question of the evolutionary challenge. Joyce here offers a first pass at an answer that relies on the notion of categorical imperatives, which are requirements that are supposed to hold unconditionally. He argues that

[f]eelings of "inescapable requirement" will, in certain circumstances, serve reproductive fitness more effectively than clear-headed calculations concerning desire-satisfaction, because they will enforce cooperative behavior, at the motivational level, more resolutely (Joyce 2001, 140).

Joyce is not here suggesting that a belief in categorical imperatives is sufficient for motivation, but rather that there is a "distinctive behavior-modifying value of moral concepts" insofar as "they are imbued with maximal authority" (140). He is arguing that moral concepts, such as categorical imperatives, are one of the best ways of getting individuals to cooperate in groups. Imagine a scenario where you need to choose a person to help you with an important task, say (somewhat anachronistically) building yourself a shelter. Let's also say that you only have two people to choose from: Claire and Noah. On the one hand, Claire feels that she can potentially maximize her desire-satisfaction by choosing to help you build your shelter. On the other hand, Noah feels a *categorical requirement* to help, that he is *morally required* to help you. Joyce says that Claire helps you build your shelter only because of her belief that she can maximize her desires by doing so. If that is indeed her motivation for helping you, then you may be suspicious

since she might jump ship for some other task that better maximizes her desire. Since Claire is only helping you because she is trying to better her own lot (i.e., maximize her desires), if she comes across a more fruitful opportunity to do just that, why shouldn't she go for it? Meanwhile, Noah will be there to help you regardless, because he feels categorically required to do so. In this way, Noah proves to be a better ally than Claire.

Again, the idea here is that having a belief in categorical imperatives, or at least being able to think in terms of categorical requirements, perhaps commits one to certain actions with greater certainty than do beliefs in desire-maximization for engaging in those same actions. Moral concepts (like categorical imperatives), then, are imbued with an authority that is lacking from mere desire-satisfaction. Tying this in with the argument above, Noah signals that he is the kind of person who believes in categorical requirements, and thus is likely to engage in reciprocal altruism. He then reaps the benefits of reciprocal altruism, and thus enhances his evolutionary-fitness.

2.6 Moral Properties

We can now pick up on the evolutionary challenge. Given the above evolutionary story, certain proponents of the evolutionary challenge (Joyce and Ruse among them) argue that the existence of *moral properties* has been accounted for. Natural selection has cultivated in us a sense whereby we respond to perceived moral properties, but *only* perceived properties. In having our desires supplanted by a stronger sense of requirement, namely *moral* requirement, natural selection works towards guaranteeing certain beneficial behaviours. And since these behaviours increase evolutionary fitness, it is likely that these behaviours would have been selected for (or, at least, not selected against). Thus,

the evolutionary challenge suggests that our moral sense either developed or was not selected against due to the fact that it helped secure certain behaviours – behaviours, mind you, that confer tremendous evolutionary advantage. So, when our moral sense is given this genealogical story, some authors posit that we are in a sense freed from having to commit to saying that moral properties *actually* exist. We seemingly do not need to make reference to any sort of objective fact that makes our moral judgments *true*.

In doing so, the evolutionary challenge mounts an attack on the moral realist position, which (roughly stated) says that there are “moral facts and truths that are objective in some way” (Brink 1989, 14).¹¹ In taking evolutionary biology seriously, Sharon Street (2006) says that moral realists will now have “to explain the relation between these evolutionary influences... and the independent evaluative truths that realism posits” (109). As such, all moral realists are bound to take the challenge seriously. The challenge asks: if we can give a plausible account of why we are inclined to believe that moral properties exist, need we presume anything further? Do we need to additionally suppose that moral properties *actually* exist, and that certain actions or events are *actually* true in reference to objective moral facts? Is any action, event, or behaviour *actually* good or bad, moral or immoral, etc., or have we just evolved to think so because it is incredibly beneficial to us to think in moral terms?

The moral realist will say that *some of our moral beliefs correspond to something objective and true*: what the realist means is that there are *moral facts* that make our moral beliefs true or false. Conversely, defenders of the evolutionary challenge will say

¹¹ I give a fuller treatment of moral realism in (3.1).

that we would have our moral beliefs irrespective of the existence of objective moral facts. This latter sentiment is precisely what leads Ruse to say that the belief in objectively true moral facts is an “illusion foisted upon us by our genes” (Ruse 1998, 162). He means of course that we can account for our moral beliefs without having to suppose the truth of those beliefs, and that it is only an evolutionary influence that leads us to believe in moral objectivity. As such, the evolutionary challenge attempts to undermine the possibility of moral realism.¹²

2.7 The Darwinian Dilemma

The basic thrust of the evolutionary challenge, as presented by Joyce, seems to me successful insofar as it undermines our presupposition that our moral beliefs are true. But there are some weaknesses in Joyce’s account that can be shored up. To do so, I will exposit one of the stronger versions of the evolutionary challenge: Sharon Street’s so-called *Darwinian Dilemma*. As she says, this is the problem of accounting for, on the one hand, the moral truths that realism posits, and on the other hand the moral beliefs that emerged because of evolutionary influence (Street 2006). An important qualification before diving in: the targets of Street’s argument are *all* realist theories of value (i.e., moral, epistemic, aesthetic, etc.), while my focus here is simply on *moral* realism. As

¹² One may ask: is the utility function of having moral beliefs constitutive of an objective moral fact? That is, if evolutionary fitness is the reason why we are inclined to suppose the existence of moral properties, then couldn’t that fitness itself be identified as a real (and thus objective) moral property? Thanks to David Matheson for this point. I respond by noting that this seems to be the wrong type of objectivity to focus on. This type of objectivity does not provide any normative force necessary for robust moral claims, and thus is unlikely to be the type of objectivity that moral realists have in mind. In particular, this type of objectivity does not qualify as categorical. I discuss this further in (4.1)

such, Street's arguments cast a wider net than mine, and will have far greater implications – so, my usage of her argument is considerably narrower than the argument she herself presents.¹³

Some terminology:

[evaluative facts] facts of the form that ϕ is a normative reason to ψ ; that one should or ought to ϕ ; that ϕ is good, valuable, or worthwhile; that ϕ is morally right or wrong; and so on.

[evaluative attitudes/judgments] states such as desires; attitudes of approval and disapproval; unreflective evaluative tendencies such as the tendency to experience ϕ as counting in favour of or demanding ψ ; and so on.

One of the targets of Street's *Darwinian Dilemma* is a certain form of moral realism. In particular, her argument applies to those realists who think that evaluative facts hold independently of evaluative attitudes. Of course there are finer distinctions to be made here, since the precise constitution of ϕ is what entails one's specific type of moral realism. Broadly speaking, if you think that ϕ is a *natural fact*, then you are a *naturalist moral realist*, and if you think that ϕ is a *non-natural fact*, then you are a *non-naturalist moral realist*. As I mentioned in (1.2), my focus here on non-naturalist moral realists.¹⁴

¹³ For example, David Copp (2008) offers a critique of Street (2006), though Copp looks only at moral realism. It is important to keep in mind that the success or failure of Street's evaluative anti-realism argument cannot come from looking at one subset of evaluative realism. There may be interesting ways in which evaluative anti-realism is necessary to guarantee the success of, say, moral anti-realism. If so, then my usage of Street's argument may be incomplete. In any event I do not take (what I view as) the success of Street's critique of moral realism as declarative of the success of all value realism critiques. This allows me to avoid the implication that my argument here will commit me to, say, epistemic anti-realism.

¹⁴ My focus on non-naturalism is one reason why my argument cannot completely undermine moral realism. Even if non-naturalist moral realism is unsuccessful, naturalist moral realism (or constructivist realism) may be a plausible alternative.

Since moral realists adhere to the thesis that evaluative facts hold irrespective of evaluative attitudes or evaluative judgments, we should look at some evaluative judgments. Street offers some examples (2006, 115):

- a) The fact that something would promote one's survival is a reason to favour it.
- b) The fact that something would promote the interests of a family member is a reason to do it.
- c) We have greater obligations to help our own children than we do to help complete strangers.
- d) The fact that someone has treated one well is a reason to treat that person well in return.
- e) The fact that someone is altruistic is a reason to admire, praise, and reward that person.
- f) The fact that someone has done one deliberate harm is reason to shun that person or seek his or her punishment.

What is initially striking here is the universality of these evaluative judgments. Indeed, while some moral philosophers play up the intransigence of evaluative disagreement, there does seem to be an incredible similarity amongst our most basic evaluative judgments. Still, as Street notes, evolutionary biology offers a powerful explanation for why evaluative judgments like (a)-(f) are universally present: it is because these evaluative judgments can easily be construed as contributing greatly to evolutionary fitness. To see why this is so, Street asks us to consider the converse of the above evaluative judgments (2006, 116):

- a*) The fact that something would promote one's survival is a reason against it.
- b*) The fact that something would promote the interests of a family member is a reason not to do it.
- c*) We have greater obligations to help complete strangers than we do to help our own children.
- d*) The fact that someone has treated one well is a reason to that that individual harm in return.
- e*) The fact that someone is altruistic is a reason to dislike, condemn, and punish him or her.
- f*) The fact that someone has done one deliberate harm is a reason to seek out that person's company and reward him or her.

Again, what is striking here is the rarity of evaluative judgments like (a*)-(f*). Plausibly, this is because having these latter evaluative judgments would be a severe hindrance to evolutionary fitness, and thus would likely have been selected against, or at least certainly not selected for.

Combined, the evaluative judgments which we do see, such as (a)-(f), and the evaluative judgments that we do not see, such as (a*)-(f*), serve as premises in an argument in support of Street's *Darwinian Dilemma*. This argument looks something like the following:

- i) If evolution were a powerful influence on our evaluative judgments, then we would see evaluative judgments that have considerable evolutionary advantage.
- ii) We *do see* evaluative judgments that have a considerable evolutionary advantage – i.e., (a)-(f).
- iii) We *do not see* evaluative judgments that have a considerable evolutionary disadvantage – i.e., (a*)-(f*).
- iv) Therefore evolution was a powerful influence on our evaluative judgments.¹⁵

If we found that we routinely had evaluative judgments that were evolutionarily disadvantageous, such as (a*)-(f*), then it would “constitute powerful evidence that the content of our evaluative judgments had not been greatly influenced by Darwinian selective pressures” (Street 2006, 116). But (a*)-(f*) are not what we find, and (a)-(f) are what we find, so we can plausibly conclude that evolution has greatly shaped our evaluative judgments.

This strikes me as the strongest version of the evolutionary challenge in the literature. For one, Street's version mostly avoids the problematic aspects of Joyce's. Street does not talk about psychological motivation, nor about whether desire-

¹⁵ Street presents something like this argument that she calls “the adaptive-link account” (2006, 127).

maximization is less reliable than a belief in categorical requirements. And secondly, Street's *Darwinian Dilemma* clearly presents the problem for moral realists. That is, given that evolutionary pressures have greatly informed the content of our evaluative judgments, need we bother supposing that our evaluative judgments here are true? It is Street's version of the evolutionary challenge that I work with, and will attempt to shore up given the responses from moral realists.

2.8 Raising and Addressing Berker's Worry

Selim Berker points out that Street's argument, as presented above, "leaves it mysterious why we didn't evolve merely to have the relevant motivations on their own, without any accompanying normative judgments" (unpublished, 3). The suggestion here is that it is at least conceptually possible, and perhaps evolutionarily plausible, that we could lack evaluative judgments altogether. If all that matters is that we act a certain way – i.e., in ways favourable to our survival – why do we need evaluative judgments at all? Berker's worry is essentially this: why do we have moral beliefs, as opposed to simply having the motivation for particular behaviours themselves? Berker suggests that Street's account does not readily connect moral beliefs with particular behaviours.

To be fair, Berker does not reject the possibility that there *may be* an evolutionary explanation for why we make certain evaluative judgments, nor that something like Street's account could play a role therein. His point, rather, is that a "more complete evolutionary explanation is going to lack the beguiling simplicity" that makes Street's account seem so compelling (unpublished, 3). I think that Berker points out an important worry with Street's argument, and with evolutionary challenges in general. The

connection between evaluative judgments, on the one hand, and the behaviours or actions such judgments are supposed to elicit, on the other, has not been clearly articulated above. Indeed, I noted this as one reason for preferring Street's account to Joyce's, largely because Joyce's motivational account is somewhat problematic. So, I will attempt to fill in the motivational gap in Street's account, without undoing its beguiling simplicity. And in doing so, I will not rely on the knotty motivational account that Joyce offers.

Certain of our value judgments appear to contribute to the kind of psychological profile that makes reciprocal altruism an actuality. What I mean by this is that, if we had value judgments such as (b*) or (c*), then it is quite unlikely that the behaviours typified by reciprocal altruism would take place. Fortunate for us, then, that (for whatever reason) we tend to have make value judgments like (b) or (c). Having value judgments similar to (a) through (f) is immensely beneficial, given that it facilitates fitness enhancing behaviours that arise because of reciprocal altruism.

Imagine if this were not the case. Rather than unreflectively thinking that

d) The fact that someone has treated one well is a reason to treat that person well in return.

we were instead inclined to unreflectively think that

d*) The fact that someone has treated one well is a reason to do that that individual harm in return.

If we were inclined to think that (d*) rather than (d), it is unlikely that we would engage in reciprocal altruism. Having evaluative attitudes like (a*)-(f*) does not seem to confer any sort of benefit onto the person that has them, whereas having evaluative attitudes like (a)-(f) quite clearly does. If we did have evaluative attitudes like (a*)-(f*), then it seems

that we would be significantly less likely to engage in reciprocal altruism, and thus we would not be privy to the benefits that arise because of reciprocal altruism. We can cash out this discussion as follows:

- i) Evaluative attitudes like (a)-(f) are indicative of a moral sense.
- ii) A moral sense makes it likely that we will engage in reciprocal altruism.
- iii) Reciprocal altruism proves to be fitness enhancing.
- iv) Fitness enhancement is evolutionarily advantageous.
- v) So, evaluative attitudes like (a)-(f) can be seen as evolutionary advantageous.

I do not see these entailments holding necessarily. One could have a moral sense without necessarily engaging in reciprocal altruism; one can engage in reciprocal altruism without necessarily being conferred fitness enhancement; and so on. Still, the argument I present here should ease Berker's worry since it more readily connects particular evaluative judgments with certain motivations. The action takes place at (i) and (ii): certain evaluative judgments indicate that we have a moral sense, and our moral sense gives us a disposition to engage in reciprocal altruism. And, importantly, I think that even with the introduction of these claims into Street's account, the argument still seems to be beguilingly simple.

2.9 The Basic Moral Realist Response

Nearly all moral realists who engage with the evolutionary challenge, in support or against it, seem to accept the above argument. However, in order to preserve their realism, moral realists will make additional claims. Derek Parfit (2011), for one, suggests that

our cognitive abilities, we can assume, were partly produced by evolutionary forces. But these abilities later ceased to be governed by these forces, and had their own effects. Natural selection gave us wings, but when we could fly, we

soared into the sky. We used these cognitive abilities to discover some new kinds of truths (520).

Similarly, Thomas Nagel (2001) suggests that stumbling upon these *new kinds of truths*

is like stepping into what looks like a small windowless hut and finding oneself suddenly in the middle of a vast landscape stretching endlessly out to the horizon (72).

As I read it, this simile presents the idea that evolution has led us to have certain beliefs, but once we start examining those beliefs we may find that we have gained access to *other* beliefs that are not evolutionarily influenced. Evolution leads us to the hut, but once inside we find that the inside is much larger than we had imagined. In essence, moral realists like Nagel and Parfit accept that some of our evaluative attitudes are evolutionarily influenced, but suggest that other evaluative attitudes are not similarly influenced. They are here suggesting that there is no evolutionary advantage to *certain* evaluative moral judgments, and as such we cannot interpret our having such judgments as evidence of their evolutionary advantage. Furthermore, they will claim that the reason we have these latter judgments is because they are *true*.¹⁶

In Chapter 3 I will give a more thorough analysis of Nagel's and Parfit's responses, and in Chapter 4 I will show how the general structure of such responses seem to further undermine the possibility of moral realism. For now, I will note simply that even though most moral realists tend to accept that some our evaluative judgments are strongly influenced by evolution, some moral realists do not take this fact to undermine

¹⁶ Cf. Parfit. Concerning the beliefs that 'pain is bad' or 'we have strong reasons to promote our own survival and the survival of our children': "these beliefs are *not* badly mistaken, but correspond to some of the independent normative truths. Pain *is* bad, and we *do* have strong reasons to promote the survival and well-being of ourselves and our children" (Parfit 2011, 532).

the possibility of moral realism as a metaethical position. The unstated premise here is that our moral judgments can be separated into two types: basic and advanced. Some moral realists seem to suggest that our basic judgments are evolutionarily influenced, while our advanced judgments are not. So far, defenders of the evolutionary challenge seemed to have ignored this duality. For my part, while I do not think that moral realists have preserved moral realism, I do think they have implicitly noted an important distinction that needs to be seriously explored. As said, in subsequent chapters I develop this implicit suggestion from moral realists, and show why, contrary to their intentions, it in fact only serves to further undermine the possibility of moral realism.

2.10 Summary

In this chapter I outlined an evolutionary account of our moral sense, presented the evolutionary challenge to moral realism, and introduced the notion of moral realism. I explicated what I take to currently be the strongest version of the evolutionary challenge (i.e., Sharon Street's *Darwinian Dilemma*) and bolstered it given the criticisms of Selim Berker. I then introduced the general framework for how moral realists have responded to this challenge.

These steps setup the next chapter where I will give a more robust definition of moral realism and carve out a specific version of that metaethical position that I think is best targeted by the evolutionary challenge. Following that, I will flesh out the moral realist response to the evolutionary challenge from both a non-naturalist perspective and a theistic one. In doing so I will highlight the weakness of each of those responses. This will allow me to, in the final chapter, unveil the general argument form that moral realists

are relying on. I will show that no matter which particular argument moral realists use, whenever they use the specific argument form that separates moral beliefs into basic and advanced types, the possibility of moral realism is further undermined.

Chapter 3: Why Should We Be Moral Realists?

3.1 A Quick Look at Moral Realism

First, let us look at a few definitions of moral realism, so that we can focus on the target of the evolutionary challenge. David Brink says that moral realism posits the existence of moral facts and truths that are objective in some way (Brink 1989, 14).

Russ Shafer-Landau suggests that moral realism

is the theory that moral judgements enjoy a special sort of objectivity: such judgements, when true, are so independently of what any human being, anywhere, in any circumstance whatever, thinks of them (Shafer-Landau 2005, 2).

Geoffrey Sayre-McCord says that moral realists think that

moral claims do purport to report facts and are true if they get the facts right. Moreover, [realists] hold, at least some moral claims actually are true (Sayre-McCord 2001).

And Mark Timmons states that moral realists believe

that there are moral facts (or properties) that are independent of human attitudes, conventions and the like (Timmons 1999, 106).

What these definitions have in common is the idea that moral realism suggests that moral facts are objective and mind-independent. I accept these two characteristics, and later defend them (in 4.1 and 4.3, respectively). Still, the definitions I have selected here will undoubtedly prove unpleasing or incomplete to some self-described moral realists; indeed, there seem to be as many definitions of moral realism as there are moral realists. My goal is not to conclusively settle the definition of moral realism but simply to target one of its most common versions, i.e., the one relayed by Brink, Shafer-Landau, Sayre-McCord, Timmons, and others. To summarize the above accounts, and to echo the definition from Street in the last chapter, if someone thinks that the moral fact that ϕ

holds independently of anyone's evaluative attitude or judgments, then that person is most likely a moral realist. In what follows I presume such an account of moral realism.

3.2 An Additional Reason for Moral Realism

Recall from (2.9) that both Parfit and Nagel accept that some of our basic evaluative attitudes have been evolutionarily influenced. Yet, in defending moral realism, both suggest that though some of our beliefs cannot be said to necessarily track truth, we have other beliefs that are not evolutionarily influenced, and thus these latter beliefs may still be candidates for realism. This is the basic structure of the moral realist response to the evolutionary challenge: when faced with the fact that some beliefs cannot be said to necessarily track truth, argue that we may nevertheless have additional beliefs that are not similarly suspect.

There are multiple ways of defending moral realism in this way. In this chapter the strategy that I will focus on involves the positing of an *additional reason* to be a moral realist. In addition to Parfit and Nagel, Alvin Plantinga presents a similar argument: all posit that, though some of our beliefs may have been evolutionarily influenced, we have an additional reason to think that some of our beliefs are true. Parfit and Nagel argue that the additional reason to presume that some of our moral beliefs are true is our rationality itself, and Plantinga argues that God ensures the truth of some of our moral beliefs. Like Parfit's and Nagel's response, I argue that Plantinga's response is wanting, and thus moral realism still faces severe difficulties as a plausible metaethical position. I start here with a deeper analysis of Nagel and Parfit.

3.3 The Platonists¹⁷

[I]f, per impossible, we came to believe that our capacity for objective theory were the product of natural selection, that would warrant serious skepticism about its results beyond a very limited and familiar range. An evolutionary explanation of our theorizing faculty would provide absolutely no confirmation of its capacity to get at the truth. Something else must be going on if the process is really taking us toward a truer and more detached understanding of the world.¹⁸

- Thomas Nagel (1986, 79)

If we were disposed to have certain normative beliefs because these beliefs were advantageous, and these beliefs would have been advantageous whether or not they were true, we would have been disposed to have these beliefs whether or not they were true. Natural selection would have been a distorting influence, since this cause of these beliefs would have been unrelated to their truth. That would give us reasons to doubt that these beliefs are true.

- Derek Parfit (2011, 513)

Parfit and Nagel argue that, even though some of our moral beliefs cannot necessarily be said to track truth, we have an additional reason to suppose that other of our moral beliefs are in fact true. The additional reason here is the (supposed) notion that human rationality allows us to access necessarily and objectively true facts. Parfit and Nagel say that, much like human beings rely on our rationality to know that $2+2=4$, we can rely on our rationality to know that some actions are morally permissible and that others are not.

Parfit (2011) goes so far as to suggest that

¹⁷ It may be debatable whether Parfit is actually a platonist. In his (2011) he does not self-describe as a platonist, but rather as a *non-metaphysical non-naturalist cognitivist*. It is unclear to me whether the difference between platonism and Parfit's alternative are substantial. However, supposing there is indeed an important difference between them, we can say that Parfit avows those features of platonism that are incompatible with metaphysical naturalism.

¹⁸ Note Nagel's idiosyncratic usage of *per impossible* reasoning here. Typically such reasoning is employed for patently absurd counterfactuals like "if, per impossible, 3 were divisible to 2, it would an even number" or "if, per impossible, circles had corners..." etc. The fact that Nagel places the evolutionary origin of our ability to reason on this level of absurdity might be why he doesn't take these arguments as seriously as he should.

just as cheetahs were selected for their speed, and giraffes were selected for their long necks, human beings were selected for their rationality. That may be how we became able to reason validly, and respond to reasons (494).

And Nagel (2012) suggests that

[R]easoning [is] correct or incorrect in virtue of something independent of the thinker's beliefs, and even independent of the community of thinkers to which he belongs. We take ourselves to have the capacity to form true beliefs about the world around us, about the timeless domains of logic and mathematics, and about the right thing to do. [We] assume that there is a real world, that many questions, both factual and practical, have correct answers, and that there are norms of thought which, if we follow them, will tend to lead us toward the correct answers to those questions. [We] assume that to follow those norms is to respond correctly to values or reasons that we apprehend. Mathematics, science, and ethics are built on such norms (71).¹⁹

The suggestion here is that our rationality allows us to see (in some non-perceptual way) that, say, $2+2=4$ is true. Parfit says that “even if we cannot yet explain how we came to have this ability [i.e., rationality], we can justifiably believe that we can recognize such necessary truths” (2011, 520).

Both Parfit and Nagel attempt to associate the apparent success of mathematics with the hope of success in the moral domain. They argue that just as $2+2=4$ stands as an objective truth which we have discovered by reason, so too are certain moral propositions objectively true and discoverable by that same (or an approximately similar) cognitive faculty.²⁰ In effect, this strategy sees moral realism looking up admirably at the apparent

¹⁹ Curiously, throughout *Mind and Cosmos* (2012), Nagel often adds aesthetics to the list of objectively true domains.

²⁰ Neither Parfit nor Nagel are very explicit about the cognitive faculties at work in our apparent recognition of objectively and necessarily true mathematical and moral facts. Instead, they issue vague statements about how even if we don't know how this faculty works, it nevertheless *does* work.

success and authority of mathematics, in the hopes that some of those qualities might trickle down. I call this, somewhat pejoratively, *the piggybacking strategy*.²¹

I am suspicious of this strategy, of using success in mathematics to argue for the possibility of success in moral theorizing. My suspicions are twofold. First, in (3.4) I make use of what I call an *external argument*, whereby I discount the supposedly objective and necessary truths of mathematical and moral claims by undermining the ontology these claims are set within.²² The effect here is that neither mathematical nor moral claims can be called necessarily and objectively true in the way required to justify realism in either domain. Effectively, then, there is no success in mathematics that morality may piggyback on. My second strategy, in (3.5), is what I call an *internal argument*. Here I question the specifics of Parfit's and Nagel's arguments about how we are thought to know that there non-natural moral facts. They argue that since we can respond to *moral reasons* we thereby have indirect evidence that *non-natural moral facts* exist. I unpack this argument and show how Parfit and Nagel leave it mysterious why we should think moral reasons come from non-natural moral facts.

²¹ Others might want to call such strategies *companions in guilt arguments*. Cf. Cuneo (2007) and Lillehammer (2002). This latter phrasing is not terribly clear to me. After all, what exactly are the so-called 'companions' supposed to be guilty of? Piggybacking seems to me the better word choice, since one domain (i.e., morality) gains from the strength or authority of another domain (i.e., mathematics).

²² Jonas Olson uses a similar strategy, though he doesn't talk about mathematical claims. With respect to moral and epistemic claims, Olson distinguishes between *categorical* and *imminent norms* and *categorical* and *hypothetical reasons*. Categorical norms and reasons hold categorically, while imminent norms and hypothetical reasons are contextualized to certain (agent-specific) ends. See Olson's (2011).

3.4 Mathematics and Morality

While the intuition that $2+2$ actually equals 4 is quite strong, and more bedrock than any moral claim, it seems that there are a few ways to undercut this particular mathematical claim. I do so here, with admittedly *prima facie* implausibility, by denying that $2+2$ actually equals 4 in an objective and mind-independent way. I start by adding a prefix operator to the mathematical claim here. Indeed, my strategy for dealing with mathematical claims, in general, is simply to prefix any supposedly objective mathematical truths with a bracketed or context-specific operator. While in an untutored sense we may think that

$$2+2=4$$

I think we are only justified in claiming that

$$[\text{According to our mathematical axioms}] 2+2=4.$$

By “our mathematical axioms” I mean the tenets of mathematics that we currently accept. The further question of whether those axioms are themselves true is, I think, a question without an answer. We can say that some of our mathematical statements are true, but only when we implicitly prefix them by saying “according to our mathematical axioms”, or “following from accepted axioms”, or something similar. In the same way that certain moves in chess are correct or incorrect based on accepted axioms that outline permissible chess moves, so are certain mathematical claims true or false by reference to accepted mathematical axioms. I say more about this account of mathematics below.

One thing that follows from prefixing mathematical claims in this way is that *the entire mathematical claim*, prefix and all, becomes objectively true. While I do not think it is objectively true that $2+2=4$ *simpliciter*, I do think it is objectively true that, according

to our mathematical axioms, $2+2=4$. That is, second-order claims about what is or is not true according to a certain axioms are justifiable, but first-order unprefix claims are not. Again, this response has to overcome its *prima facie* implausibility, an implausibility that perhaps stems from the difficulty in imagining a counterfactual where $2+2$ did not actually equal 4. So to justify my response a brief digression into the ontology of mathematics is necessary.

If we claim that $2+2$ actually equals 4, we are *platonists* about mathematics.

Mathematical platonism makes the following two claims (Balaguer 2011):

- i) there exist abstract mathematical objects (i.e., non-spatiotemporal mathematical objects), and
- ii) our mathematical sentences and theories provide true descriptions of such objects.

So a mathematical platonist will say that

- i*) there exist abstract mathematical objects 2, 4, +, =, and
- ii*) our mathematical sentence " $2+2=4$ " provides a true description of those objects.

Even though (i*) and (ii*) are very intuitive, note that mathematical platonists make a very strong claim: they are saying that 2, 4, and so on are objects that *actually exist*. It is not as if, when pressed, the platonist will say that by *actually exist* they simply mean that we merely use or work with such objects. Rather, what the mathematical platonist means is that such objects actually exist in a non-spatiotemporal manner.²³ Thus, if one is a mathematical platonist, they are most certainly *not* a metaphysical naturalist – which is, again, the view that only natural things exist. If one thinks that objects like 2 and 4

²³ Furthermore, mathematical platonists are not just mentalists about mathematical objects. That is, platonists do not simply believe that "2", "4", and so simply exist in our minds.

actually exist, then one is committed to saying that non-natural objects exist in some yet undefined sense.

But mathematical platonism is not the only option for making sense of our mathematical claims. *Fictionalists* about mathematics deny that mathematical objects actually exist in the platonist sense, and thus fictionalists preserve the possibility of metaphysical naturalism. I believe we should be mathematical fictionalists, and not platonists. Furthermore, I think that mathematical fictionalism results from a denial of platonism, so I will here offer reasons to reject platonism and default to fictionalism.²⁴

Earlier I mentioned the so-called *Benacerraf Problem*, which is the worry about how human beings can have knowledge of non-spatiotemporal objects (see the footnote in 1.3). Hartry Field restates the problem:

Benacerraf's challenge – or at least, the challenge which his paper suggests to me – is to provide an account of the mechanisms that explain how our beliefs about [mathematical objects] can so well reflect the facts about them. The idea is that *if it appears in principle impossible to explain this*, then that tends to *undermine* the belief in mathematical entities, *despite whatever reason we might have for believing in them* (Field 1989, 26).²⁵

Field goes on to say that

we should view with suspicion any claim to know facts about a certain domain if we believe it impossible in principle to explain the reliability of our beliefs about that domain (Field 1989, 233).

²⁴ Despite Hartry Field's attempts at nominalization (1980) and (1991), it seems as though we simply must use mathematical statements in our best science. As such we cannot simply be error theorists about mathematical objects. Given that we have no good reason to accept platonism about mathematical objects, fictionalism seems to me our only alternative. For positive accounts of mathematical fictionalism, see Stephen Yablo's (2002) and (2005).

²⁵ Italics in original.

Platonists should certainly worry about the Benacerraf Problem, but it need not undermine the *possibility* of mathematical objects. After all, the platonist could respond to the problem like this: *Sure, we don't yet know how we can have knowledge of mathematical objects, but clearly we can utter true mathematical statements, like "2+2=4". The best explanation about how we can do this is that mathematical objects like 2, 4, +, and = actually exist. So even if we can't currently explain our mathematical knowledge, perhaps one day we will. In the meantime, our true mathematical statements assure us that an answer will be (or could be) forthcoming.*²⁶ What Field suggests, however, is that platonists are not justified in presuming that an answer to the Benacerraf Problem even *could be* forthcoming. He argues that not only do we *currently not* know how knowledge of mathematical objects is possible, such knowledge is *in principle* impossible. Field writes that "mathematical [objects] don't seem to be causally involved in producing physical effects" (Field 2001, 329). Plainly, we do not need to suppose that mathematical objects like "2" and "4" exist in order to make sense of our having the belief that "2+2=4". And even if we did need to suppose that 2 and 4 exist in order to make sense of our beliefs, 2 and 4 seem incapable to producing those beliefs. While it is relatively uncontroversial that natural objects causally affect what we observe (and thus influence our beliefs), there does not seem to be any way *in principle* that 2 and 4 causally affect *anything*, nor any way to make sense of something non-natural affecting something natural. So there seems to be no way to make sense of mathematical objects affecting human beings.

²⁶ A mathematical platonist arguing along these lines might say that presuming that mathematical objects actually exist is *indispensable* to our best mathematical theories.

Some versions of mathematical fictionalism avoid these problems by focusing less on the existence mathematical *objects* and more on *what we can do with* mathematics. Field, for one, points out how we obviously

don't need to make mathematics actually be about anything for it to be possible to objectively assess the logical relations between mathematical premises and mathematical conclusions (Field, 2001: 317).

On this view, mathematical sentences are not true by reference to mathematical objects, but simply as a matter of being "a *consequence* of accepted axioms" (Field 2001, 320).²⁷ This is but one fictionalist strategy. I do not explore positive accounts in detail here, but what fictionalist accounts of mathematics have in common is the suggestion that we do not have to move from having the belief that $2+2=4$ to thinking that 2 and 4 are objects that actually exist. And when we are not forced to say that non-natural or non-spatiotemporal objects exist, we maintain our prior commitment to metaphysical naturalism.

This account of mathematical ontology is sufficient for understanding the piggybacking argument employed by Parfit and Nagel. The argument is this:

- i) Mathematical objects exist.
- ii) Some of our mathematical sentences provide true descriptions of mathematical objects.
- iii) If moral facts exist, they are like mathematical objects.
- iv) When moral sentences are true they provide true descriptions of moral facts.
- v) Some moral sentences are true.
- vi) So moral facts exist.

I think that we can easily undermine this piggybacking strategy. If we are metaphysical naturalists, we should say that mathematical objects would be metaphysically strange

²⁷ Italics in original.

(i.e., non-spatiotemporal and/or non-natural). This gives us a reason to deny (i). The denial of (i) then gives us an exclusive disjunction for (ii): either our mathematical sentences are false, or our mathematical sentences are true but not by reference to mathematical objects. I think we should say that our mathematical sentences are true insofar as they follow from accepted axioms. Therefore, we should say that the truth of our mathematical sentences does not give us a reason to think that mathematical objects exist. We can simply appeal to the usefulness of our mathematical axioms: accepting certain tenets allows us to make useful mathematical claims.

Both Parfit and Nagel explicitly deny metaphysical naturalism.²⁸ Presumably, it is their acceptance of mathematical platonism that motivates their acceptance of moral realism, and it is easy to see why this is so.²⁹ If one can readily accept that non-spatiotemporal mathematical objects actually exist, then one will have an easier time believing that non-natural moral facts actually exist. Once one bites the bullet on mathematical platonism (or something very close to it), one has already abandoned metaphysical naturalism, and it thereby becomes easier to accept that non-natural moral facts might exist. Given the intuition that mathematical claims imply the existence of mathematical objects, one might be motivated to abandon metaphysical naturalism. And it does not seem to be moral realism that leads one to abandon naturalism, since the

²⁸ Cf. Parfit (2001, 502) for the so-called *Validity Argument* where he expressly denies metaphysical naturalism.

²⁹ Again, Parfit denies that his view of mathematics is platonist: see Parfit (2011: Chapters 31-32) for his so-called “Non-metaphysical Non-naturalist Cognitivism” alternative to platonism. For my purposes, Parfit’s metaphysical position is sufficiently close to platonism insofar as what makes certain mathematical claims true is something non-natural.

intuitions concerning non-natural moral facts are not in themselves as strong as intuitions concerning non-spatiotemporal mathematical objects. Hence, moral realism here piggybacks on the work done by mathematical platonism. If mathematics motivates us to abandon metaphysical naturalism, non-naturalist moral realism instantly becomes plausible.

I have said that Parfit and Nagel offer us no good reason to abandon metaphysical naturalism. But suppose that we are mathematical platonists and thus accept (i), must we then accept (iii)? I think not. The piggybacking strategy is by no means an obvious way of motivating us to accept moral realism. Combined, (i) and (ii) suggest that true mathematical statements reference non-spatiotemporal mathematical objects, but why should that give us a reason to think that moral statements thereby reference non-natural moral facts? Presumably Parfit and Nagel assume that since moral statements are true, they must be true in reference to *something*. But as Field (2001) shows, we can have mathematical objectivity *without* supposing that mathematical objects exist: our mathematical statements could be true “simply [as] a matter of... being derivable from explicitly or implicitly accepted axioms” (317). If so, the piggybacking argument fails: we can make moral claims without supposing that non-natural moral facts exist. Our moral claims could be true merely as a matter of following from accepted moral axioms.³⁰

So Parfit and Nagel have not presented a convincing argument to abandon metaphysical naturalism, nor a good reason to think that morality parallels mathematics

³⁰ And, again, the further question of whether those axioms themselves are true does not seem truth-apt.

in the way necessary for realism. Indeed, if we are metaphysical naturalists, the parallel we should draw between mathematics and morality is that moral claims (like mathematical claims) are not true descriptions of actually existing non-natural facts (or objects). Instead, moral and mathematical claims are “correct” insofar as they follow from accepted axioms, and we cannot even ask whether the axioms *themselves* are true (since the truth or falsity of an axiom cannot be proven). As such, the *prima facie* implausibility of denying that $2+2=4$ should now seem justified, bracketed to an acceptance of metaphysical naturalism.

3.5 Social and Cultural Influence

My second objection to Parfit and Nagel concerns the way in which they posit the existence of non-natural moral facts. They are committed to saying that abstract objects actually exist, but how can we come to know about such objects given that they are non-natural or non-spatiotemporal? As said, this is the well-known Benacerraf Problem. The issue is essential this: how can human beings, who are spatiotemporal beings, have access to non-spatiotemporal objects? Again, how can we gain access or have knowledge of metaphysically mysterious objects?

Parfit and Nagel offer a work-around solution. Rather than accessing these objects or facts directly, we can prove their existence indirectly. This solution relies, again, on human rationality. The argument is this: if human beings can use reason to access mathematical truths, then we can also use reason to access moral truths. In particular, human beings can come to recognize that we can respond to *reasons*, and there are reasons to believe that, for example, some of our moral beliefs are true. If moral facts

exist, these facts are non-natural are so cannot be directly accessed (*per* the Benacerraf Problem). But human beings are rational agents and can respond to reasons. Moral facts, so it is suggested, provide us with moral reasons. And since we can respond to moral reasons we thereby have indirect evidence of moral facts.

Presented in this way, the argument does seem to get around the Benacerraf Problem. If there is an unresolved issue concerning how human beings can have knowledge of non-spatiotemporal objects or non-natural facts, then the above argument provides a solution. Parfit and Nagel will say that we do not have direct access to such objects or fact, but rather that such objects or facts provide *reasons*. And since human beings can have contact with reasons, we have indirect evidence of non-spatiotemporal objects or non-natural facts.

Though plausible, I do not think that this work-around argument is convincing. We can ask: what exactly is going on when we respond to a moral reason? Do we have to say that a moral reason has necessarily been provided by a non-natural moral fact? What guarantee do we have that we are actually responding to non-natural moral facts, as opposed to reasons provided by *some other* fact? While it might be possible that we may be responding to reasons such facts, it seems to me just as likely (indeed, *more* likely) that we are responding to reasons provided by our currently accepted moral axioms. Let me spell out this objection.

When we think that we are responding to the

[non-natural moral fact] Helping the less fortunate is morally good.

when we have the belief or think to ourselves that

We have reason to help those less fortunate than ourselves

in fact what we might really be responding to is that

[According to our moral axioms] Helping the less fortunate is morally good.

The prefix operator “according to our moral axioms” cashes out here in the same way that “according to our mathematical axioms” did earlier. That is, our moral axioms are simply the currently accepted moral tenets of our society. While the moral realist will say that moral facts are best understood as an axiom that is isomorphic with reality, anti-realists will deny that isomorphism and posit alternative origins. Robert Nozick (2001), for one, asks:

Might not our intuitions be due largely (or wholly) to shaping by societal forces and to cultural conditioning? If so, would not the resulting ethical principles purportedly justified on the basis of the intuitions be as contingent and variable as the particular social factors that produce such intuitions? (237)

If we agree with Nozick, we should say that particular societies moralize certain actions or behaviours in certain ways. Our moral intuitions, then, are true in reference to accepted moral axioms. Alternatively, relying on the account of morality presented in Chapter 2, we might say that particular actions or behaviours are moralized because of the fitness-enhancing benefits conferred on those who do so. Whatever the case may be, it does not seem obvious that simply because we can respond to reasons we must thereby be responding to reasons stemming from non-natural moral facts.

Plausibly, we can say that reasons do indeed come from facts, however not all facts provide reasons. If we have a reason to ϕ there is likely some fact or set of facts that provides us with our reason to ϕ . If we take ourselves to have a moral reason to, say, donate to charity, then there is likely some fact or set of facts that provides us with that reason. But if we do not assume that there are non-natural moral facts, then the mere fact that we take ourselves to have a moral reason to donate to charity fails to have the

explanatory power that Parfit and Nagel think it does. As I see it, there are at least two scenarios:

Scenario A

Fact 1: objectively, donating to charity is morally good.

Moral Reason: we have a reason to donate to charity.

Scenario B

Fact 2: according to our moral axioms, donating to charity is morally good.

Moral Reason: we have a reason to donate to charity.

In Scenario A, *Moral Reason* is provided by *Fact 1*. In Scenario B, *Moral Reason* can be provided by *Fact 2*. On this latter interpretation, what we take to be a moral reason comes not from something categorical, but from societal forces and cultural conditions. Even though the moral reason in Scenario's A and B are identical, the *source* of that reason is importantly different. I see no reason to think that moral reasons can only come from *Fact 1*, while it is uncontroversial that moral reasons can come from *Fact 2*. The fact that we take many historical moral codes to have been mistaken is evidence enough that we think morality *can* be social or cultural in nature. The challenge is thus for the moral realist is to prove that morality is not *just* social or cultural. Still, Scenario B is less problematic than Scenario A, because Scenario A makes a superfluous metaphysical claim about facts (i.e., that non-natural facts exist), as well as an unjustified claim about our ability to access such facts (*per* the Benacerraf Problem).³¹

³¹ A moral realist might agree that we do have *some* moral reasons that are provided by societal forces and cultural influence, but press that there are *other* sources of moral reasons. This is an unattractive position, and I critique it in the next chapter.

To sum up: first, in presuming that moral realism can piggyback on the apparent success of mathematics, Parfit and Nagel presume the truth of mathematical platonism. But if we are metaphysical naturalists we ought to reject all forms of platonism. If we are not persuaded to abandon metaphysical naturalism, then paralleling morality with mathematics does not seem to accomplish what Parfit and Nagel hope it will. And second, in attempting to solve the problem of how we can access non-natural facts, Parfit and Nagel do not provide any assurance that moral reasons can only come from such moral facts. As I argued, it is more plausible that our reasons come from social or cultural sources.

3.6 The Theist

To this point I have been focusing on how some moral realists have posited that, in light of the evolutionary challenge, we have an additional reason to presume that moral realism is true. Parfit and Nagel suggest that rationality assures us that our moral beliefs can be true. Alvin Plantinga takes a different approach: he is a theist, and he uses his belief in God to motivate his commitment to moral realism. For Plantinga, the reason we can presume that our moral beliefs are true is because God ensures that they are. But analyzing Plantinga's theism is not my primary focus; I have nothing specific to say about theism beyond noting again my acceptance of metaphysical naturalism. Here I will be focusing on Plantinga's suggestion that there is an irresolvable tension for the moral realist who wishes to accept evolutionary theory: one can either be a moral realist, or accept evolutionary theory, but not both.

Plantinga argues that the above tension between evolutionary theory and moral realism generalizes to *all* of our cognitive faculties, and not just to our moral beliefs. If evolutionary considerations give us reason to doubt whether our moral beliefs are true, then perhaps evolutionary considerations also give us reason to doubt our memory, perception, and so on. According to Plantinga, unless we posit an additional reason (beyond evolutionary pressures) to think these cognitive faculties are reliable, then evolutionary theory not only undermines our moral sense providing us with true beliefs, but any of our cognitive faculties providing us with true beliefs. For Plantinga, the supernatural is what ultimately grounds all of these faculties: our memories, perceptions, intuitions, etc., are trustworthy because of God.

Plantinga frames the issue by saying that evolutionary considerations appear to undermine the possibility of robustly true moral beliefs. He argues that if we think “our cognitive faculties have been cobbled together by natural selection” then we cannot sensibly believe that such faculties are reliable (Plantinga, 2012, 313). What he means here is that we have no grounds for assuming they are *necessarily* reliable. Plantinga's argument in favour of this scepticism is very much the same one I presented in Chapter 2 – that is, evolutionary pressures do not necessarily track truth, rather they track evolutionary fitness. Thus, if our cognitive faculties are the result of a non-truth-tracking process, then we should not believe that these faculties necessarily produce true beliefs. Unless, of course, we have an additional reason for doing so. Plantinga suggests that metaphysical naturalists who accept evolutionary theory lack an additional reason for

realism, and thus cannot coherently be moral realists. If natural selection need not produce true beliefs then it almost certainly does not produce true beliefs.³²

Plantinga's argument is essentially this: a commitment to metaphysical naturalism should undermine our trust in all of our cognitive faculties. Evolutionary influence works towards fitness-enhancement, and not necessarily truth. In effect this is the evolutionary challenge to moral realism spread out to encompass all of our cognitive faculties. For Plantinga, the metaphysical naturalist should not only be suspicious of our moral sense's ability to provide true beliefs, but equally suspicious of *any* cognitive faculty providing us with true beliefs. Where metaphysical naturalism turns on evolutionary theory, evolutionary influence provides a defeater for the "intuitive assumption that [our] cognitive faculties are reliable" (Plantinga 2011, 314). Metaphysical naturalists, then, have no reason to think that any of their beliefs are true.

Even though Parfit and Nagel are *not* metaphysical naturalists, we can apply this argument to them. As we saw, Parfit and Nagel argue that, in spite of the undermining influence of evolution on our moral sense, rationality nevertheless assures us that we can have true moral beliefs. Plantinga would presumably respond to Parfit and Nagel as follows: if the evolutionary origin of our moral sense gives us reason to suspect that it can provide us with true beliefs, then we cannot legitimately use *another* evolutionarily influenced faculty to make up the difference. That is, if our moral sense on its own cannot

³² He argues that the likelihood of us having true beliefs is exceedingly low. Plantinga (2011) writes: P(R/N&E) is low, where "R" is the proposition that our cognitive faculties are reliable, 'N' is naturalism, and 'E' is the proposition that we and our cognitive faculties have come to be in the way proposed by the contemporary scientific theory of evolution. 'P(.../___)' is shorthand for 'the probability of... given ___' (309). See Plantinga's (2011: Chapter 10) for his *evolutionary argument against naturalism*.

assure us we have true moral beliefs, then it is equally true that our rationality on its own cannot help assure us that our moral beliefs are true, nor that *any* of our beliefs are true.³³ So where Parfit and Nagel suggest that we can know that some moral claims are true because we are rational, Plantinga would argue that they are relying on yet another non-truth-tracking faculty to reach that conclusion.

We are by now familiar with this general argument strategy so I will not spend time on it here. Instead, I will do two things: first, I respond to Plantinga's suggestion that the scepticism arising from a commitment to metaphysical naturalism (and evolutionary theory) generalizes to *all* cognitive faculties. I argue that he is mistaken to think that moral scepticism qua evolutionary theory generalizes to all cognition. And second, I address the suggestion that we do in fact have an additional reason to suppose that our moral beliefs are true (i.e., because God ensures that they are). I conclude that Plantinga's additional reason to be a moral realist is no better than Parfit's or Nagel's, and metaphysical naturalists should not be persuaded.

3.7 Medium-sized Dry Goods in our Immediate Vicinity

I think we can get around the problematic wide-generalizability of Plantinga's argument by appealing to *the indispensability of certain perceptions*. When we observe (to use J.L. Austin's phrase) medium-sized dry goods in our immediate vicinity, the reliability of our beliefs about those objects is I think secured by their indispensability. To the question, *is there actually a computer screen in front of you, or have you evolved to think there is?*, I believe we can answer that both are true – we evolved to think there is such an object in

³³ Street makes a similar point in her (2006, 124).

front of us, and that there *actually is* such an object in front of us. In this way, the content of our beliefs is very closely linked with what is true. I think we should say that our perceptual faculties evolved to be sensitive to true information from our environment. Of course, we may sometimes get false positives here (due to, say, shadows, mirages, hallucinations, etc.), though in general our beliefs concerning whether or not there are medium-sized dry goods in our immediate vicinity are trustworthy.

There are at least two good reasons to think so. First, it is difficult to make sense of the idea that most of our beliefs about objects in our immediate vicinity are in fact false. While I think we can give a plausible account for why we are inclined to think that there are moral facts even though none likely exist, I know of no similarly plausible story for why we would continually think that there are objects in our immediate vicinity if there were not in fact such objects there.³⁴ In this way, our beliefs about medium-sized dry goods in our immediate vicinity are indispensable in a way that our moral beliefs are not. The role that the former beliefs play is much different than the role of our moral beliefs. Though moral beliefs are surely useful for allowing us to get along with others (as argued in 2.2), and beneficial to us evolutionarily (as argued in 2.4), these are the only

³⁴ One might argue that there really are no medium-sized objects in our immediate vicinity. Appealing to quantum physics, one might say that all apparent medium-sized objects are really just subatomic particles, and so in our most reflective moments we should not truly distinguish between the particles comprising the supposed medium-sized objects and all other particles. To argue along these lines sounds like mereological nihilism, and if it is, it does not trouble my position here. I could rephrase and say that there are particles arranged computer-wise, or medium-sized-object-wise. I do not have to commit to computers or medium-sized objects strictly speaking existing. Cf. Peter van Inwagen's (1995).

features that we need to acknowledge our moral beliefs tracking.³⁵ It does not seem necessary (or desirable, given naturalism) to say that our moral beliefs reference moral facts. So we still need an additional reason to suppose that our moral beliefs are *true*, rather than merely fitness-enhancing. On the other hand, our beliefs about medium-sized dry goods seem to *only* plausibly track the existence of such objects. While we could say that it is evolutionarily useful to know whether there are objects in front of us, we should say that our beliefs about medium-sized dry goods track the existence of such objects *themselves*. Any advantage that subsequently arises is because such objects really exist. Our beliefs about medium-sized dry goods are only advantageous because there actually are medium-sized dry goods.

Another reason for denying Plantinga's claim of wide-generalizability is that we have indubitable confirmation about certain cognitive faculties, whereas we have strong and persistent disagreement concerning certain moral beliefs. By every metric we can think of, we can confirm that there is *actually*, say, a computer screen in front of you. We run into such objects (quite literally), they are physically there, and we can test whether or not they physically there. We have universal agreement here. There are no persons who, looking over your shoulder, would deny that you are looking at a computer screen. Our best science confirms the existence of such objects. And so on. If we had such

³⁵ Here I am appealing to a parsimony argument much like Gilbert Harman's (1977). In his famous example, he argues that there is no reason to assume anything about "moral facts" when we see some hooligans set a cat on fire (7). He says that "[i]t seems to be completely irrelevant to our explanation whether our intuitive immediate judgment is true or false" (7). For Harman, moral facts are unnecessary for a complete explanation of our moral judgments, and I argue so is the question of truth.

reliable confirmation concerning our moral beliefs then that would be a *prima facie* reason to suppose that they too are plausibly true.

3.8 Changing Moral Beliefs

So Plantinga is mistaken in supposing that the scepticism that undermines our moral sense providing true beliefs spreads to all cognitive faculties. I've offered two reasons to think that some of our other cognitive faculties produce true beliefs. But if Plantinga's argument does not generalize, and it only applies to our moral beliefs, what then? Do we have an additional reason to suppose that any of our moral beliefs are true? If we are metaphysical naturalists, then Plantinga says we do not have an additional reason that we can appeal to. Given that our moral sense has been cobbled together by evolution, it plausibly tracks only fitness. The metaphysical naturalist does not have any recourse to go beyond fitness-tracking: they seem unable to posit that our moral beliefs can be true. On this, Plantinga and I agree: naturalists should not accept non-naturalist moral realism. But Plantinga is not a metaphysical naturalist, and he has an additional reason to work with. His additional reason, as said, is supernatural: the God supposed by Abrahamic religions. It is this claim that I here evaluate.

In what way does God constitute an additional reason to be a moral realist?

Plantinga argues that God serves as a guarantor for our moral beliefs by providing us with a reason to think that the world is knowable. Combined with certain interpretations of God's nature (i.e., omnibenevolence, omnipotence, etc.), the fact that (for Plantinga) God exists is the reason that we can think that our moral beliefs are true. The scant evidence offered that Plantinga offers here is purely theological. He says that

[i]t's an essential part of theistic religion – at any rate Christian theistic religion – to think of God as providentially governing the world in such a way as to provide... stability and regularity (2011, 272).

He adds that

Christian theism involves the idea that God governs the world; that what happens does not come about by chance, but by virtue of God's providential governance (2011, 272).

When it was noted how our moral sense has been cobbled together by evolution, this was said to undercut the thought that our moral beliefs are true. The reason, again, is that (i) we need only have fitness-enhancing beliefs, and (ii) we have no reason to think that our moral beliefs are true. But if we take stock of Plantinga's theological commitments, we can see why he thinks our moral beliefs *can* be considered true. With the background idea that God governs the world in a reliable way, a theist has a reason to think that our beliefs are reliable. Rather than it being a stunning coincidence that we have come to have true moral beliefs issuing from an evolutionarily influenced moral sense, a theist can say that we have the moral beliefs we do because of God fixes our moral beliefs. Just as God "sets forth or promulgates laws of nature" God also "sets forth moral laws" (Plantinga 2011, 273).³⁶ As such, it is no coincidence that we have true moral beliefs: since God guides evolution and sets the moral law, the two have been correlated in order to allow us to have true moral beliefs.

Of course, non-theists are unlikely to be persuaded by this argument, least of non-theist metaphysical naturalists. The account Plantinga gives is not simply a God-of-the-

³⁶ Also, even if Plantinga acknowledges that our moral beliefs arise solely because of evolutionary influence, this would not prove problematic for him. He argues that evolution is not a blind process, but is guided by God (2011: Chapters 1 and 2).

gaps argument, where a supernatural entity stands in for a (currently) unexplained phenomenon. Rather, Plantinga *starts* with the premise that God exist, and argues that his theism provides more explanatory power than any competing theory. Indeed, it seems to be Nagel's and Parfit's argument that should be considered gap arguments: they are forced into accepting the *prima facie* implausible position of metaphysical platonism because of their assurance that certain moral (and mathematical) statements *must* be true. So on the one hand this shows the somewhat *post hoc* nature of Nagel's and Parfit's suggestion – after all, wouldn't the better move simply be to abandon moral realism rather than presume metaphysical platonism? On the other hand, because Plantinga starts with the premise that God exists, his argument is more difficult to argue against. That is, it is difficult for a non-theist to argue on Plantinga's terms, or what we can call an internal argument.³⁷

But there is at least one internal argument against Plantinga's account: how can he account for changing moral beliefs? If, as Plantinga says, God guarantees the truth of our moral beliefs, why have those beliefs changed? If God both guides evolution and sets forth the moral law, then why has there (historically at least) been a mismatch between our moral intuitions and the moral law? If we take Plantinga's account seriously, we should think that our moral beliefs and the moral law have always been perfectly coordinated. But moral intuitions concerning (for example) slavery have changed, from a prevailing attitude of permissibility to widespread condemnation. Moral realists like Nagel and Parfit have an easy explanation for this shift: while the non-natural moral facts

³⁷ Conversely, an *external* argument against Plantinga would be where one, say, argues *against* the existence of God. I have no interest in doing that here.

are necessary and eternal, we (human beings) have simply been mistaken about what they are. On this account, our changing moral beliefs have been the real-time working out of what exactly those true moral facts are.

But it doesn't seem that Plantinga can appeal to a similar answer. The somewhat intrusive sense (we might say *micromanaging way*) in which he thinks God guarantees our moral beliefs should not admit of mistakes or miscorrelation: after all, the any misalignment between our moral beliefs and the true moral facts would have to be God's fault, and one would think that this is incompatible with God's nature. If (for Plantinga) God is supposed to provide assurance that our moral beliefs are true, the fact that our moral beliefs have changed seems problematic. Without a fuller account for the exact way that God is thought to fix our moral beliefs, we have little reason to suppose that our moral beliefs are true. And where God is supposed to serve as an additional reason for moral realism, the explanation is wanting. Like the posited additional reason from Parfit and Nagel, Plantinga's solution is unsatisfying.

3.9 The Similarity between Moral Realism and Theism

To close this chapter I'll draw a comparison between the Plantinga's commitments and those of Parfit and Nagel. Obviously Plantinga's ontological commitments are articles of faith. In his (2000) and (2011), Plantinga argues that belief in God stems from a special and peculiar kind of mental faculty – a *sensus divinitatis*. Non-theists seemingly lack this faculty, thus their inability or refusal to believe in God. Belief in God, then, is not a logical entailment, but something issuing from an altogether different faculty, something which (says Plantinga) need not be subsumed by conflicting evidence from other

faculties. His often used example is of a person on trial who truly knows that they did not commit the crime in question. Even if a wealth of compelling evidence is presented that they points towards them committing the crime, Plantinga thinks that the accused is nevertheless justified in believing themselves innocent. For Plantinga, belief in God is similar: even though, as Allan Gibbard says, “with each passing decade we get further indications that naturalistically is the way we must learn to understand ourselves” (Gibbard 2012, 24), Plantinga still thinks that one is justified in believing in a supernatural being, like God. Where his *sensus divinitatis* assures him that God truly exists, Plantinga claims that he needn’t have this belief cut down by any conflicting evidence.

But what about Parfit and Nagel, indeed all non-naturalist moral realists? How do they reconcile their belief in non-natural moral facts with the fact that, with each decade we get more and more confirmation that naturalism is surely the way to understand ourselves and the universe? It seems to me that non-naturalists implicitly rely on a claim that is not unlike Plantinga’s *sensus divinitatis* claim. If implicitly, non-naturalist moral realists suggest that we are really, truly justified in thinking that certain actions, events, or behaviours can be morally good or bad. Even in the face of conflicting evidence, for example metaphysical naturalism (which assures us that non-natural things do not exist), these moral realists still say that we can justifiably believe that some of our moral beliefs are true because non-natural things simply must exist. As with the *sensus divinitatis*, no conflicting evidence is able to cut down the moral realist belief that there really are some things are *actually* moral or immoral.

Nagel is particularly forthcoming concerning this. In his (2012), he explicitly states that his acceptance of moral realism is incompatible with our naturalistic understanding of ourselves and the universe. But, curiously, when he recognizes the tension between, on the one hand, metaphysical naturalism and evolutionary theory, and on the other hand non-naturalist moral realism, Nagel rebuffs naturalism and evolutionary theory. He says:

I remain convinced that pain is really bad, and not just something we hate, and that pleasure is really good, and not just something we like. That is just how they glaringly seem to me, however hard I try to imagine the contrary... On the Darwinian account, this must be regarded as an illusion—perhaps an illusion of objectivity that is itself the product of natural selection because of its contribution to reproductive fitness... However, in my case the scientific credentials of Darwinism... are not enough to dislodge the immediate conviction that objectivity is not an illusion with respect to basic judgments of value (2012, 110)

For Nagel, because non-naturalist moral realism is incompatible with evolutionary theory and metaphysical naturalism, something has to go. Nagel confidently denies evolutionary theory and metaphysical naturalism. More accurately, throughout his (2012) Nagel denies metaphysical naturalism and argues that evolutionary theory is woefully incomplete. He argues that evolutionary theory needs to make reference to (as yet undefined) teleological principles that make the existence of moral truths inevitable. Because Nagel is confident that moral realism is true, he suggests that our metaphysical theories must account for the inevitable existence of non-natural facts.³⁸

I arrive at the opposite conclusion from Nagel. If there is an unresolvable tension between moral realism, on the one hand, and metaphysical naturalism and evolutionary

³⁸ Similarly, Nagel (2012) says that both evolutionary theory and our best metaphysics needs to account for the realist status and inevitability of consciousness, mathematical truths, and aesthetic values.

theory, on the other, then moral realism is the dispensable theory. I fail to see why we should be motivated to upend evolutionary theory and metaphysical naturalism in order to accommodate certain moral intuitions, especially since we have a plausible account of why we have those moral intuitions. While I think that Nagel is correct in identifying the tension inherent to simultaneously believing naturalism, evolution, and moral realism, I think he chooses the more problematic way of easing it. Still, if all non-naturalist moral realists were as forthcoming as Nagel, and as ready to admit that we need a radically different view of the cosmos in order to accommodate realism, perhaps fewer people (read: moral philosophers) would be moral non-naturalists.

So it should seem now that non-naturalist moral realism requires one to accept a very peculiar view of the cosmos. We might even say that this metaethical position borders on the religious or the supernatural. As Sharon Street writes:

[Moral] realism has become a strange form of religion—a religion stripped clean of everything except the bare conviction that there are independent [moral] truths that one is capable of recognizing (Forthcoming, 23).

Moral realists like Parfit and Nagel need to construct an odd metaphysical worldview in order to make sense of some of our moral intuitions. And while Plantinga doesn't posit God in order to account for his moral intuitions, he too has an odd metaphysics. What these three authors seem to have in common is an unwavering adherence to certain notions: for Parfit and Nagel, it is just glaringly obvious that some things are categorically moral and others categorically immoral; for Plantinga, it is just obviously true that God exists. All are unshakeable here, and thus construct their worldviews around these beliefs. The fact that, as Gibbard says, investigation into the universe and our place within it tells us that we should understand the world naturalistically has little

traction for Parfit, Nagel, and Plantinga. And while Plantinga acknowledges that his beliefs stem from a *sensus divinitatis*, I suspect that Parfit and Nagel will less readily admit that their beliefs seem to be, as Street says, a strange form of religion.

3.10 Summary

In this chapter I looked at two different moral realists positions. Both attempted to provide a justification for moral realism: Parfit and Nagel relied on rationality, and Plantinga appealed to God. None of these specific attempts managed to justify moral realism, nor gave us any plausible grounds for thinking that our moral beliefs are true. In the final chapter, I detail how the general moral realist response has been to separate our moral beliefs into two types: basic and advanced. This general strategy is novel, and it has thus far remained unexplored by anti-realists who support the evolutionary challenge to moral realism. Contrary to their intentions, in responding to the evolutionary challenge, moral realists have effectively undermined their preferred metaethical theory. Thus, even though Parfit's and Nagel's and Plantinga's attempts prove wanting, we should not expect that a better account will be forthcoming. I argue that no matter the particular content that non-naturalist moral realists use in responding to the evolutionary challenge, if a particular basic strategy is employed the argument will prove ultimately ineffective.

Chapter 4: Critiquing Moral Realist Responses

In this chapter I look at three strategies for preserving moral realism in light of the evolutionary challenge. The first two strategies are idiosyncratic readings of moral realism that attempt to eliminate certain problematic characteristics of this metaethical position. One, from Stephen Finlay (2008), denies that morality involves objectivity. Another, from Terence Cuneo (2007), denies that morality involves a mind-independence constraint. The third strategy denies that our moral beliefs are of a single type. This line of argument has not been explicitly stated in the literature, so in the chapter I will construct that argument and subsequently show it to be inadequate.

4.1 Morality is Objective

As I understand them, moral judgments are held as objective judgments. By this I mean that both lay and philosophical usages of moral concept suggest categoricalness. We can understand categoricalness in the standard philosophical way: a categorical imperative, say, “represent[s] action as objectively necessary of itself, without reference to another end” and “as *in itself* good” (Kant [1785] 1996, 67). However, Stephen Finlay (2008), for one, argues that our moral concepts and judgments simply do not presuppose categoricalness.³⁹ If Finlay is right then the anti-realist is mistaken in focusing so much on categoricalness. And if Finlay can strike categoricalness as a necessary characteristic of moral realism, this metaethical position instantly becomes more palatable.⁴⁰ To block

³⁹ For replies to Finlay that differ from my own, see Jonas Olson’s (2010), and Richard Joyce’s (2011).

⁴⁰ I do not outline Finlay’s full objection here, though to summarize he argues that moral claims are taken to be relational rather than absolute. See his (2008).

this objection, I offer three compelling reasons to think our moral concepts do indeed involve objectivity:

- 1) It accounts for our inclination to distinguish between social conventions and moral rules.
- 2) It accounts for typical features of our moral phenomenology.
- 3) It accords with the empirical evidence we have about the psychology of meta-ethics.

I outline these reasons here.

First, following Elliot Turiel (1983) it seems clear that we make a distinction between conventional rules and moral rules, and to accept a conceptual distinction between the two. It is commonly held that conventional rules depend on some kind of authority figure – be that a specific person, an institutional norm, etc. When school children raise their hands before speaking they are following a conventional rule. Conventional rules are somewhat arbitrary, in that people are inclined to think they are changeable. If, one day, school children are told that they no longer need to raise their hands before speaking, the children will probably follow this new rule. On the other hand, moral rules are thought to be obligatory, universally applicable, and impersonal (Turiel 1983, 36). Unlike conventional rules, moral rules are typically held to be non-arbitrary, in that they cannot simply be changed on a whim. If school children are told that they may slap other children whom they disagree with, the children will not adopt this new rule. Indeed, children are able to draw a distinction between conventional and moral rules within their first few years (Smetana and Braeges 1990). There is much more that can be said about the moral-conventional distinction, but for my purposes here I wish simply to highlight the fact that moral rules are categorical in a way that conventional

rules are not. The fact that moral concepts involve categoricalness supports this distinction.⁴¹

Second, the phenomenology of moral experience, at least as I experience it, points towards our moral concepts being imbued with significant objectivity. Moral objectivity is more prevalent in negative moral concepts than in positive moral concepts. For example, those who employ the concept of something being *morally reprehensible* seem to feel that the concept holds categorically – i.e., it's not just that one *just feels* or *just thinks* that, say, harming children for fun is morally reprehensible, rather it is that one believes that harming children for fun *actually is* morally reprehensible. I suspect that those who employ this kind of moral concept have difficulty making sense of those who disagree with them: if you think that harming children for fun is morally reprehensible, whereas someone else does not, you're unlikely to entertain the notion that you simply have different opinions. I suspect that you are more likely to think that that other person is just plain wrong.⁴² Though there are likely to be exceptions, I think most people are inclined to say that harming children for fun *just is* morally reprehensible, and that's that.⁴³

⁴¹ Though I accept Turiel's distinction, see David Shoemaker's (2011) and Daniel Kelly, et als. (2007) for potential problems with the moral/conventional distinction.

⁴² You might also be inclined to think that the person who does not think that harming children for fun is morally reprehensible is *themselves* morally reprehensible simply for their thinking so.

⁴³ Contrary to the literature on moral dumbfounding (Cf. Haidt [2012]), I think that people will be able to provide some reasons for their strong moral views. However, I suspect that these people will, in turn, take the reason they provide to hold categorically.

As said, I think such strong views seem more forthcoming for moral prohibitions than they do for moral recommendations – that is, moral condemnation seems to invoke more steadfast stances than moral commendation. For example, when we wield the concept of *moral praise*, the objectivity or categoricalness perhaps slackens. When it comes to moral praise, I think here we may have an easier time accepting the fact that others may disagree with us. If you think that saving children from a burning building is a morally praiseworthy action, whereas someone else does not think so, I think you're less likely to think that that other person is just plain wrong. You may be more willing to entertain the idea that the two of you simply have different opinions about what is in fact morally praiseworthy. These intuition pumps get at the categoricalness that seems embedded in certain of our moral concepts; it seems to be moral condemnations that are thought to be more categorical, as compared to moral recommendations.

So much for armchair reflection. The phenomenology of moral experience is an empirical issue and could be studied as such – namely, how *do* people actually interpret or understand their usage of moral concepts? The third reason I have for thinking that our moral concepts involve a degree of objectivity comes from the psychology of metaethics. In Goodwin and Darley's (2008) study participants were presented with four types of statements: factual, ethical, social conventional, and taste. Examples:

[factual] Boston is further north than Los Angeles.

[ethical] Robbing a bank to pay for an expensive holiday is a morally bad action.

[social conventional] Wearing a bath robe to a seminar is wrong behaviour.

[taste] Frank Sinatra was a better singer than Michael Bolton.

First, participants were asked to regard these types of statements as either true, false, or as an opinion/attitude. Second, participants were told that the experimenters had a body of data from prior testing, and that none of the statements produced complete agreement or

disagreement amongst those tested. The participants were then asked to interpret disagreement with their own attitudes. For example, if one judged that *Robbing a bank in order to pay for an expensive holiday is a morally bad action* was true (or false, or an opinion/attitude) how would they then interpret someone who disagreed with their judgment? Four options were available here:

- i) The other person is surely mistaken.
- ii) It is possible that neither you nor the other person is mistaken.
- iii) It could be that you are mistaken, and the other person is correct.
- iv) Other.

Goodwin and Darley note that, in general, factual statements were treated as more objective than ethical statements, ethical statements were treated as more objective than statements of social convention, and statements of social convention were treated as more objective than statements of taste. This suggests a range of objectivity, with factual statements at one end, and taste statements at the other. Ethical statements, it seems, lie closer to factual statements. As the authors note:

For ethical statements, the most common response (50 out of 100) was fully objective [followed by] intermediately objective (28 out of 100; 11 of these were “true (or) false, but neither party need be mistaken”; and 17 were “opinion, but the other party is surely mistaken”) (Goodwin and Darley 2008, 1348).

However, the authors of the study also noted that “unlike philosophers of meta-ethics lay individuals are objectivists about some ethical beliefs but not others” (1358). Whether or not one judged an ethical statement as fully objective was “highly sensitive to the content of the claims in question (i.e., robbery vs. abortion), and not merely to whether the claims are generally agreeable” (1346).

I take this study as showing a statistical likelihood that non-philosophers interpret their ethical statements as more objective than, say, matters of convention or taste. If

there is a range of objectivity, ethical statements appear to sit closer to statements of fact than to convention. Still, one may disagree with these conclusions, perhaps by suggesting that those who participated in these studies were not competent users of moral concepts. So the objection might go: when competent concept users (presumably moral philosophers) are questioned, different results may emerge. I leave this debate to others, and proceed with strong evidence in favour of our ordinary moral concepts having a notion of categoricalness inherent to them.

4.2 Is Morality Mind-independent?

Throughout the preceding chapters I have made many suggestions that moral realism involves so-called objective and mind-independent moral facts. In the literature this is called the *mind-independence constraint*. I see adherence to the mind-independence claim contributing to the implausibility of moral realism as a metaethical position because of my acceptance of metaphysical naturalism (1.3). In particular, having mind-independent moral properties in one's ontology seems to me a form of metaphysical platonism, which any metaphysical naturalist is likely to find disagreeable. What proves problematic is that the exact nature of these mind-independent entities is left mysterious, but also that such entities lack any explanatory power. I have suggested that we can have a full understanding of morality without bloating our ontology with mind-independent entities.

While I take a mind-independence constraint to be implied by the most prominent versions of moral realism (3.1), Terence Cuneo (2007), for one, questions whether this is indeed the case. So here I will raise Cuneo's concerns, and show why they are

unsuccessful or unwarranted. I conclude that a mind-independence constraint proves essential to moral realism, and thus moral realism is unpalatable for metaphysical naturalists.

We can understand the mind-independence constraint as the claim that “a realist account of moral facts implies that such facts are mind-independent” (Cuneo 2007, 45). One version of this constraint holds that moral facts are mind-independent insofar as they are independent of “the attitudes of actual or hypothetical human moral agents” (45).⁴⁴ In essence, this is the claim that even if there did not exist actual or hypothetical human agents, there would still be some things that are right or wrong, good or bad, moral or immoral, and so on. To motivate his denial of the mind-independent constraint Cuneo distinguishes between *predicative* and *general moral propositions*, as well as distinguishing between *particular* and *general moral facts*.

Cuneo’s treatment is dense, so a bit of unpacking here. Particular moral facts are those facts that correspond to predicative moral propositions:

[particular moral fact] This murder is wrong.

And general moral facts correspond to true general moral propositions:

[general moral fact] Murder is morally bad.

What this distinction between particular and general moral facts makes clear is that the mind-independence constraint only applies in certain cases. That is, the constraint cannot plausibly be understood as being thought to apply to particular moral facts. For example, there is no way to make sense of someone suggesting that the

⁴⁴ Cuneo calls this the “strong version of the mind-independent constraint”. He also expounds a “weak version”, which I do not address here since I argue that even the strong-version accurately applies to moral realism. For the weak version, see his (2007, 45).

[particular moral fact] that this murder is wrong is thought to be true independent of any human moral agent. Obviously, the claim *that this murder is wrong* is a particular moral fact that applies to a very specific situation – the particular moral fact here picks out a particular murder. As such, particular moral facts are necessarily dependent on the existence of some human moral agent, i.e., an agent that is picked out by, say, the particular murder in question. So we should say, then, that the mind-independence constraint plausibly applies only to general moral facts. For example, the

[general moral fact] murder is morally bad does *not* on the face of it seem to apply to *any particular* human moral agent, irrespective of any particular situational features. Rather, general moral facts seem to apply *in general*, to *all* particular moral agents. As such, the general moral fact that *murder is morally bad* is an apt candidate for the kind of claim that a moral realist will take to be made true by some mind-independent moral fact.

So on the one hand we have particular moral facts, which by their very definition only apply to some particular human moral agent in a particular circumstance. On the other hand we have general moral facts, which by their definition do not *just* apply to a particular human moral agent. They apply to *all* particular human moral agents.

How then does the mind-independent constraint apply to general moral facts? The main reason for believing that the constraint applies to general moral facts, says Cuneo, is that it is usually understood that moral realists “believe that some moral facts exist necessarily” (Cuneo 2007, 47). So, if some moral facts exist, these facts exist necessarily.

If something exists necessarily it is mind-independent. Therefore, if moral facts exist, they are mind-independent.

For his part Terence Cuneo accepts that moral facts exist necessarily, but denies that this necessity claim entails mind-independence. He argues that there are at least two ways of denying this, a *less popular* strategy and a *more popular* one:

[less popular] a modal anti-realist may claim that all modal facts existentially depend on the occurrence of contingent states of affairs in the world.

[more popular] one may claim that putatively necessary moral facts existentially depend on the attitudes of hypothetical moral agents.

Cuneo offers no examples of moral realists who subscribe to the less popular means of denying that necessity entails mind-independence, rather he merely presents it as a so-far unexplored option.⁴⁵ However, the more popular option is the route of many moral realists, in particular so-called *naturalist moral realists* (e.g., the Cornell Realists, Peter Railton, etc.).⁴⁶ A typical example of this more popular option would read as follow:

The putatively necessary fact *murder is wrong* (or ϕ) supervenes on the necessary fact *that ideally rational human agents in epistemically ideal conditions would disapprove of murder* (or ψ).

I will shorten this option to say that:

The putatively necessary moral fact ϕ supervenes on the necessary fact ψ .

So, if the more popular option is true, then it seems to dispute the mind-independence claim: what makes the putatively necessary moral fact ϕ true is the fact that it supervenes on the necessary fact ψ , and ψ (as said) is how certain hypothetical agents would respond in certain situations. If this more popular strategy is successful, then one

⁴⁵ Since it's a view that no one apparently holds, I set it aside here.

⁴⁶ Cf. Richard Boyd's (1988), Peter Railton's (1986), and Nicholas Sturgeon's (1985).

cannot say that moral beliefs or propositions are made true by some mind-independent moral fact. Instead, one would have to admit that moral beliefs or propositions could be made true by the evaluations or attitudes of hypothetical persons (who, presumably, have minds). If so, then moral realism can successfully eliminate the mind-independence constraint. And, if moral realists avail themselves of this strategy, moral realism becomes ostensibly more palatable as a metaethical position insofar as it avoids the metaphysically problematic mind-independence claim.

4.3 Preserving Mind-independence

Though I grant that a moral realist could argue along these lines, there seem to be serious worries in doing so. I raise two problems with Cuneo's treatment of the mind-independence constraint.

First, what does it mean to say that there could not exist *hypothetical* human moral agents? As Cuneo articulates the mind-independence constraint, it says that moral facts are made true by something independent of *actual* or *hypothetical* human beings. I can make sense of the claim that there could not exist actual human beings, but struggle to make sense of the claim that there could not exist hypothetical human beings. Part of my critique of moral realism is directed at the claim that, even if there were no human beings, there would still be things that are right or wrong, good or bad, etc. - and this, as said, seems metaphysically strange. However, if Cuneo is extending the mind-independence constraint to apply to both *actual* and *hypothetical* human beings, then what would it even mean to say that hypothetical human beings do not exist?

I am underwhelmed by the suggestion that, even if no actual human being thinks so, it is possible that some hypothetical human beings might find, say, ϕ morally bad. Of course, this is going to depend on a response like this:

While no actual human being finds ϕ morally bad, some ideally rational hypothetical human being in epistemically ideal conditions would find ϕ morally bad.

So while no actual person finds ϕ morally bad, if there were a perfectly rational person in a particular circumstance then that person might find ϕ morally bad. This is the often-used *ideal observer*.⁴⁷ But I do not find such thought experiments very persuasive, at least when the attempt is to argue for some objective and necessary fact. My objection follows.

One reason we might suppose that, despite the beliefs of all *actual* persons, a *hypothetical* person who is perfectly rational and in ideal epistemic conditions may find ϕ morally bad is if there were indeed some actual human being who presently thought so. That is, it seems like ideal observers are employed when we need to justify at least one actual person's currently held moral belief. If, say, Noah is the only one who thinks to himself that ϕ is morally bad, then he might give an argument for why other people should believe (as he does) that ϕ is indeed morally bad. Noah might attempt to persuade other people that ϕ is morally bad by saying to them:

If you were ideally rational and in ideal epistemic conditions, then you too would think that ϕ is morally bad.

Or maybe he might say:

Someone who is ideally rational and in ideal epistemic conditions would find ϕ morally bad, and that gives us reason to think that ϕ is indeed morally bad.

⁴⁷ Cf. Roderick Firth's (1952) and Richard Brandt's (1954).

But do we have reason to suppose that Noah has elucidated the method for finding categorical moral truths? Does it not seem more reasonable to suggest that what Noah has actually done is appealed to a hypothetical person who is (somehow) perfectly rational and in epistemically ideal conditions in order to justify what he, Noah, already believes? This latter interpretation seems correct. An appeal to an ideal observer does not seem to give us reason to think that the beliefs of *hypothetical* human moral agents can hold independent of *actual* human moral agents. Indeed, it seems to me that it is always *actual* human moral agents who put the words into the mouths of *hypothetical* human moral agents, and so what one supposes that perfectly rational persons in ideal epistemic conditions might think will always be suspiciously biased towards one's own moral convictions.⁴⁸

But supposing that Cuneo can get around that concern, I think there is a second and deeper worry with his argument. In attempting to eliminate the mind-independence constraint, it seems to me that Cuneo has simply pushed the justification for moral realism back one step. We can see why this is so by using the distinctions that he presents. Recall that the

[general moral fact] murder is morally bad

is what the mind-independence claim is thought to properly apply to. According to moral realists, what is supposed to make this general moral fact true is a *necessary* moral fact

⁴⁸ There is much more that could be said about this issue. To strengthen his claims, perhaps Cuneo should say this: *anyone* who was perfectly rational and in ideal epistemic conditions would *have to* have the same moral beliefs. Parfit suggests something like this, which he calls a "convergence claim" (Parfit 2011, 546).

that is in some way mind-independent. But Cuneo outright denies this mind-independence claim, saying instead that the

[general moral fact] murder is morally bad

could be made true by the

[putatively necessary moral fact] that ideally rational human agents in epistemically ideal conditions would disapprove of killing.

Or, as I noted earlier: the putatively necessary fact ϕ supervenes on the necessary fact ψ .

But even if the moral realist appeals to this move, we seem justified in asking two questions: first, why think that the putatively necessary fact ϕ supervenes on the necessary fact ψ ? And second, why think that the necessary fact ψ is in fact true (or necessary)?

The first question essentially asks why we should be inclined to think that, say, *murder is morally bad* supervenes on the fact *that ideally rational human agents in epistemically ideal conditions would disapprove of killing*. Why couldn't we say, tongue in cheek, that "something is morally good if and only if it is pink with yellow trimmings" (Stevenson 1937, 14)? More seriously, I do not see how any brute assertion here is supposed to give us criteria for what exactly is morally good or bad. Given the account of the pro-social origins of morality in Chapter 2, it seems as if we already have a plausible account for why we are inclined to think that *murder is immoral*, and it is an account that need not make any reference to anything being *actually* immoral.

The second question essentially asks why we should be inclined to think that ψ . That is, why think that ideally rational human agents in epistemically ideal conditions would approve or disapprove of ϕ ? The answer from moral realists is likely to be that ϕ *just is* morally bad because ψ *just is* true. So again, I would ask, what makes ψ not only

true, but necessarily true? The moral realist seems to have little to say here other than to assert that ψ *just is* true. This strategy may be successful if one can be motivated to accept that there are irreducibly true normative facts. But doesn't this mean that the mind-independence constraint, if it *does not* apply to ϕ then it certainly *does* seem to apply to ψ ? If ϕ isn't targeted by the mind-independent constraint because it existentially depends on the necessary fact ψ , then what exactly is it that makes ψ true if not some kind of mind-independent claim?

So, as I said, Cuneo seems only to have succeeded in pushing the argument back one step, and thus he has not satisfactorily done away with the claim that moral realism has inherent to it a mind-independence constraint. Even if one granted that ϕ was not mind-independent, and that it existentially depended on the necessary fact ψ , it would still seem that whatever it is that makes ψ true is indeed mind-independent. If so, then, contrary to Cuneo, it would seem that moral realism cannot actually eliminate the mind-independence constraint. As such, moral realism is indeed committed to the mind-independence constraint, and all the spooky metaphysical commitments that come along with it.

4.4 Basic and Advanced Moral Beliefs

As we saw in Chapter 3, there have been attempts to preserve moral realism in light of the fact that our moral sense has almost certainly been cobbled together by evolution. Because our moral sense originated in this way, it then becomes difficult to believe that our moral beliefs are true. To preserve moral realism, the realist needs to make up some ground. Realists need an answer for why we should think that our moral beliefs are true.

Parfit and Nagel suggest that we can come to know that our moral beliefs are true because of our rationality. Plantinga uses theism: God guides evolution in order for us to come to recognize the moral law God promulgates. I showed why both of these strategies are wanting.

But I think there is an even deeper worry. While one may look at the attempts from Parfit, Nagel, and Plantinga and see errors, one might nevertheless think a better account will be forthcoming. That is, even though recognizing the evolutionary origin of our moral beliefs appears to undermine the belief that our moral intuitions are true, we might think it is still possible that we have *other* moral beliefs free from that influence. While the evolutionary influence on our basic moral beliefs is well established, perhaps not *all* of our moral beliefs are of this type. It seems to me that realists are saying that we can distinguish between *basic* moral beliefs and *advanced* moral beliefs.

When Parfit says that

our cognitive abilities, we can assume, were partly produced by evolutionary forces. But these abilities later ceased to be governed by these forces, and had their own effects... We used these cognitive abilities to discover some new kinds of truths (Parfit, 520)

I think he might have something like the distinction between basic and advanced beliefs in mind. Others have gestured at similar distinction. Christopher Toner writes that, though our moral sense was the result of natural selection, it has been “co-opted... for purposes other than natural selection” (Toner 2011, 529). Kevin Brosnan notes that “if our moral beliefs are subject to a tracking failure, it does not follow that they are probably false” – indeed, there may be other reasons to suppose that they are true (Brosnan 2010, 55). And Fritz Alhoff acknowledges the strength of the evolutionary challenge, but suggests that “contract-based approaches to morality can be sensitive to

those presuppositions while still vindicating morality” (Alhoff 2009, 97). As I read them, what each of these authors is suggesting is that, though some of our moral beliefs are almost certainly false, we might have other moral beliefs that are not.

Okay, realist says, I can agree that some of our beliefs are suspect, but we have other beliefs that are not. This is an intuitive response, and it is one that has yet to be addressed by defenders of the evolutionary challenge to moral realism. To rectify this oversight, I will here follow up on the suggestion that we have at least two types of moral beliefs. On the one hand we have moral beliefs which have almost certainly been selected for because of evolutionarily forces. If so, it is highly unlikely that these moral beliefs are made true by objective and mind-independent moral facts. And on the other hand we may have beliefs which have may not been evolutionarily influenced, at least not in any obvious or trivial sense. If so, these other beliefs may yet be candidates for beliefs made true by the moral facts. Again, though our basic moral beliefs are likely not true, we might say that our advanced beliefs could be.

Recall the evolutionary challenge to moral realism: natural selection tracks fitness, not necessarily truth, and if our moral sense stems from natural selection then we can undermine the belief that our moral beliefs are true. But how exactly is the evolutionary challenge supposed to undermine moral realism *writ large*? At the very least it should make us question our confidence in our basic moral beliefs, i.e., the beliefs we have because they proved evolutionarily advantageous. But why suppose that those are the only moral beliefs we have? It seems that there has been an oversight here: anti-realists have jumped from talking about *some* of our moral beliefs to making a larger claim about *all* of our moral beliefs. So the anti-realist says: because our basic moral

beliefs are probably false, we have no reason to suppose that any of our moral beliefs are true. But surely this move is too quick. As said, we are inclined to make the following judgment:

b) The fact that something would promote the interests of a family member is a reason to do it.

And we are inclined to make this judgment because of evolutionary pressures. But does it follow that *all* of our moral judgments are the result of similar pressures? Not in an obvious way. Consider the following moral belief:

g) The fact that something would promote the interests of future generations is a reason to do it.

An evaluative judgment like (g) does not have an obvious evolutionary advantage, in the way that (a)-(f) almost certainly do.⁴⁹ And if not, a moral realist might then be inclined to say that (g) is not a basic moral judgment at all, but rather it is one of our advanced moral judgments. As such, the undermining effect that evolution has on our basic moral beliefs might not apply to our advanced ones. Moral realism, then, might be salvageable. In the final part of this chapter I examine what follows from this response.

4.5 Problems with the Moral Realist Response

If they avail themselves of the above strategy, moral realists now seem to be in the following position. On the one hand they readily acknowledge that our basic beliefs are almost certainly greatly influenced by natural selection – and if so, they are all probably false. On the other hand, moral realists claim that some of our advanced beliefs might be

⁴⁹ That's not to say that (g) does not have any evolutionary advantage, but rather that whatever advantage (g) has is not as superficial as (a)-(f).

separate from this influence, and thus are true. Splitting up evaluative judgments into basic and advanced types seems to preserve the possibility of moral realism. But I think that moral realists who argue in this way find themselves in an unenviable position. By splitting our evaluative beliefs into those which are basic and those which are advanced, it brings along a few odd conclusions:

- 1) The possibility that some of our advanced beliefs may be true does nothing to confer truth onto our basic beliefs
- 2) We cannot say that any of our basic evaluative judgments are true. Since basic judgments are distorted by evolution, we cannot say that we have them because they are true – they track fitness rather than truth.
- 3) It is plausible to construe many (if not all) of our advanced beliefs as comprised of basic beliefs.

I expand on these points here.

4.6 False Hope for Moral Realism

Even if our advanced beliefs were made true by some objective moral facts, this would not on its own guarantee that our basic moral beliefs are true. Consider the following conversation:

Noah: ϕ is morally good.

Claire: I don't think so. We are inclined to think ϕ is morally good because thinking that way confers tremendous evolutionary advantage on the people who have those beliefs. So our moral beliefs track fitness, not truth. We shouldn't say that our moral beliefs are true.

Noah: Ah, but not *all* of our moral beliefs track fitness. Take, for example, the belief that ψ is morally bad. If we do not have the belief because that ψ is morally bad because it was fitness enhancing, then we might have it simply because this belief is true.

Claire: True, we *might* have that belief because it is true - but what does that have to do with the belief that ϕ is morally good? We still lack a reason to think that this latter belief is true, regardless of whether or not the belief that ψ is morally good turns out to true.

To date, the conversation between moral realists and defenders of the evolutionary challenge has only developed to Noah's second reply, the conditional claim. Presented with the evolutionary challenge to moral realism, moral realists have rightly responded by saying that it is not obvious that *all* of our beliefs exist because of evolutionary influence. And if we have some of our moral beliefs because of some *other* reason, then we might be justified in thinking that we have some of our moral beliefs because they are true.

But the conversation can easily continue on from here. Why think that the mere possibility that we have certain moral beliefs absent any fitness-enhancing benefits does anything to prevent our most basic moral beliefs from being undermined? Plainly, the fact that *ψ is morally bad* might be a true moral belief does nothing to prevent *ϕ is morally good* from being a suspect moral belief. In much the same way that anti-realists have moved too quickly from the fact that *some* of our moral beliefs are likely false to the pronouncement that *all* moral beliefs are false, so too have certain moral realists moved quickly from the fact that *not all* moral beliefs are obviously false to the pronouncement that *some of our moral beliefs are true*. But this move doesn't hold: at best, I think the moral realist here is justified in holding out a hope for moral realism, and not in presuming that some of our moral beliefs are true.

4.7 Our Basic Beliefs are probably false

If the moral realist makes the move in (4.6) and holds out a hope for moral realism, then moral realism looks like it is in real trouble. By accepting that our basic moral beliefs have been evolutionarily influenced, and thus are probably false, moral realists may have

backed themselves into a corner. They would seem to be all too willing to disregard our basic evaluative judgments as false and unreliable, seemingly unbothered by or unaware of the way in which this appears to further undermine the possibility of moral realism.

What kind of robust moral theory is available if we have to eliminate or stop trusting our basic moral beliefs? For one, a basic evaluative judgment like (e) turning out to be false seems highly problematic if one wishes to maintain a commitment to moral realism. The moral realist should ask themselves what kind of realist morality they could even posit if, say

e) the fact that someone is altruistic is a reason to admire, praise, and reward that person

could not be taken as true. Given the above account, whatever else the moral realist wants to say about morality, they will now have to acknowledge that certain foundational moral beliefs, like (e), cannot be considered true any longer. We no longer have any guarantee that merely because we have a moral belief it is thereby a true moral belief. Indeed, the moral realist has already accepted that evolutionarily influenced attitudes cannot justifiably be thought to track truth, rather they track fitness. So what about (e)? When people *do* act altruistically, what can the moral realist say? How can they attempt to praise this person, or must they be silent? What kind of realist moral theory can exist if we cannot say that altruistic persons should (categorically) be praised? It seems that acknowledging that (e) is evolutionarily influenced, and thus is probably false, for is a troubling entailment for those wishing to preserve some form of moral realism.

Equally problematic is that the following judgment can no longer be considered true:

f) The fact that someone has done one deliberate harm is reason to shun that person or seek his or her punishment

Like (e), the evaluative judgment (f) is also now a suspect candidate for a true moral belief. If someone has done one deliberate harm, we no longer have objective grounds for arguing that that person should be shunned or punished. If the moral realist has already accepted that our basic moral evaluations are suspect, then a moral belief like (f) falls by the wayside. To see why (f) proves particularly problematic, we can expand the sentiment behind it. Let us say that the fact that someone has intentionally done something immoral, and thus done deliberate harm, is a reason to shun that person or seek his or her punishment. In essence, we will want to say that we have a reason to shun or punish deliberately immoral persons. But the moral realist who now accepts that (f) is non-truth-tracking cannot make or accept this expanded claim. The expanded version of (f) seems just as evolutionarily tainted as (f) itself – so if (f) is off the table, so is the expanded version of (f).

My point here is not to go through each of our basic evaluative judgments and show that the moral realist can no longer rely on their intuitive plausibility in assembling their new moral theory. Rather, I think that acknowledging that our basic evaluative judgments are probably false seems to do irreparable harm to morality as something objective and real. It now seems difficult to make sense of moral realism once basic moral judgments cannot be relied on in elucidating our moral theories.

But let us suppose that some brilliant and resourceful moral realist manages to do just this. Though our basic moral beliefs are now undermined, and thus cannot be used in our moral theories, perhaps there is a clever way to make morality simply about advanced

beliefs. If so, then the fact that our basic moral beliefs are theoretically useless might not be as disastrous as it seems. However, if this happens, I think another problem comes up.

4.8 Our Advanced Beliefs are built from our Basic Beliefs

Though it seems obvious now that our evaluative judgments should be split into basic and advanced beliefs, the relationship between these two types is not immediately apparent.

While moral realists (like Parfit, Toner, Brosnan, Alhoff, etc.) hint at a distinction between basic and advanced beliefs, it seems to me more plausible to recognize advanced beliefs as *emerging from* or *built out of* basic ones. That is, perhaps our supposed advanced beliefs are not so much *advanced* as they are *complex* – i.e., constructed from basic beliefs (or maybe we should call them *simple beliefs*). Understanding advanced beliefs in this way then seems to undermine the possibility of their realist status.

Above I mentioned one possible advanced belief:

g) The fact that something would promote the interests of future generations is a reason to do it.

How should we interpret this belief? Is it advanced, in that it is radically distinct in kind from our basic moral beliefs? This seems to be what the aforementioned moral realists want to suggest. However, an advanced belief like (g) seems to me plausibly built out of a number of basic beliefs. For example, inherent to (g) are at least the following notions:

- i) that the interest of others matters
- ii) that the fact that someone is innocent is a reason to avoid doing harm to them
- iii) that the fact that we are in a position of power over someone is a reason not abuse that power
- iv) that there is reason to think that what promotes the interests of future generations also promotes our interests
- v) that we have reason to be altruistic
- vi) that altruism can be applied to future generations

And so on. Viewed in this way, a moral belief like (g) can be seen as a complex moral belief – that is, an amalgamation of more simple moral beliefs.

But if (g) is actually composed of other beliefs, like say (i) through (vi), then why think that (g) is a *true* moral belief? More to the point, when exactly are we supposed to start talking about realism? From the above argument (4.7), we cannot justifiably think that our basic moral beliefs are true. However, many of (i) through (vi) seem to be either basic moral beliefs *themselves*, or easily broken down into basic moral beliefs. For example, there seems to be an obvious evolutionary advantage to thinking that, say, the interests of other people besides ourselves matter, or that we have reason to be altruistic, etc. So at least some of the beliefs that (g) might be comprised of are *themselves* probably false, since it is likely that they too have been greatly influenced by natural selection.

An argument along these lines, which reduces all advanced moral beliefs to basic ones, perhaps hints at some kind of nominalist project.⁵⁰ Though perhaps the moral realist can respond by saying that, even though our advanced beliefs seem to simply be composed out of basic beliefs, moral realism emerges once those basic beliefs are arranged in a certain way. I am not going to explore the plausibility and results of these strategies here, but will note simply that if the moral realist wants to preserve realism they have some considerable work to do.

⁵⁰ As far as I know, no one has attempted a nominalist project concerning our moral beliefs.

4.9 Why Be a Moral Realist?

Pointing out that we have moral beliefs that are not obviously evolutionarily influenced is insufficient to maintain moral realism. When moral realists respond to the evolutionary challenge in this way, they have not sheltered moral realism from criticism. As I've shown, the line of response which divides our moral beliefs thusly has the unintended effect of further undermining the possibility of moral realism. Contrary to their intentions, I've shown that no matter the particular content that non-naturalist moral realists uses in responding to the evolutionary challenge, if the general strategy of implicitly distinguishing between basic and advanced moral beliefs is employed, the argument will ultimately prove ineffective.

Non-naturalist moral realism, then, seems an immensely difficult metaethical position to defend. It hardly seems worth the bother, especially when there is an acceptable alternative theory on hand: moral anti-realism. What moral anti-realism lacks in terms of numbers of adherents, it more than makes up for in things like (as I've shown) accordance with metaphysical facts, confirmation from our best science, explanatory power, parsimony, and so on. While anti-realism doesn't allow us to confirm our pre-theoretic intuitions concerning morality, it might have the upshot of more readily connecting us with our normative ethical theories. If there are no non-natural facts, and thus no normative ethical positions that we objectively *have to* accept, then whatever normative ethics we *do* accept could be of our choosing. And there might be something to be said for thinking that our ethics belong to us in this way, rather than thinking they are divined or intuited from *out there* in the cosmos.

References

- Alhoff, Fritz. 2009. "The Evolution of the Moral Sentiments and the Metaphysics of Morals". *Ethical Theory and Moral Practice* 12 (1): 97-114.
- Balaguer, Mark. 2011. "Fictionalism in the Philosophy of Mathematics". *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. URL = <<http://plato.stanford.edu/archives/fall2011/entries/fictionalism-mathematics/>>.
- Bekoff, Marc. 2008. *The Emotional Lives of Animals*. New World Library.
- Bekoff, Marc and Jessica Pierce. 2010. *Wild Justice: The Moral Lives of Animals*. University of Chicago Press.
- Benacerraf, Paul. 1973. "Mathematical Truth". *Journal of Philosophy* 70 (19): 661-679.
- Berkeley, George. 1790/2009. *Principles of Human Knowledge*. Oxford University Press.
- Berker, Selim. Unpublished. "The Meta-ethical Irrelevance of Evolutionary Theory".
- Boyd, Richard. 1988. "How to be a Moral Realist". In *Moral Realism*. Edited by G. Sayre-McCord. Cornell University Press. 181-228.
- Brandt, Richard. "The Definition of an 'Ideal Observer' in Ethics". *Philosophy and Phenomenological Research* 15 (3): 407-413.
- Brink, David O. 1989. *Moral Realism and the Foundations of Ethics*. Cambridge University Press.
- Brosnan, Kevin. 2011. "Do the Evolutionary Origins of our Moral Beliefs undermine Moral Knowledge?" *Biology and Philosophy* 26 (1): 51-64.
- Carrier, Richard. 2005. *Sense and Goodness Without God: A Defense of Metaphysical Naturalism*. AuthorHouse.
- Copp, David. 2008. "Darwinian Skepticism about Moral Realism". *Philosophical Issues* 18 (1): 186-206.
- Cuneo, Terence. 2007. *The Normative Web*. Oxford University Press.
- Field, Hartry. 1980. *Science Without Numbers: A Defence of Nominalism*. Princeton University Press.
- . 1991. *Realism, Mathematics, and Modality*. Blackwell Publishing.

- , 2001. *Truth and the Absence of Fact*. Oxford University Press.
- Finlay, Stephen. 2008. "The Error in the Error Theory". *Australasian Journal of Philosophy* 86 (3): 347-369.
- Firth, Roderick. 1952. "Ethical Absolutism and the Ideal Observer". *Philosophy and Phenomenological Research* 12 (3): 317-345.
- Flack, Jessica and Frans de Waal. 2000. "Any Animal Whatever". *Journal of Consciousness Studies* 7 (1-2): 1-29.
- , 2007. "Context modules signal meaning in primate communication". *PNAS* 104 (5): 1581-1586.
- Gibbard, Allan. 2012. "Five Girls on a Rock". *London Review of Books* 34 (11): 22-24.
- Goodwin, Geoffrey P. and John M. Darley. 2008. "The Psychology of meta-ethics: Exploring Objectivity". *Cognition* 106 (3): 1339-1366.
- Haidt, Jonathan. 2012. *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. Pantheon.
- Harman, Gilbert. 1977. *The Nature of Morality: An Introduction to Ethics*. Oxford University Press.
- van Inwagen, Peter. 1995. *Material Beings*. Cornell University Press.
- Joyce, Richard. 2001. *The Myth of Morality*. Cambridge University Press.
- , 2008. "The Error in 'The Error in the Error Theory'". *Australasian Journal of Philosophy* 89 (3): 519-534.
- , 2013. "The Evolutionary Debunking of Morality". *Reason and Responsibility* 15th Edition. Edited by J. Feinberg and R. Shafer-Landau. Cengage.
- Kant, Immanuel. 1785/1996. "Groundwork of the Metaphysics of Morals". *The Cambridge Edition of the Works of Immanuel Kant*. Translated by Paul Guyer and Allen Wood. Cambridge University Press. 37-108.
- Kelly, Daniel, Stephen Stich, Kevin J. Haley, Serena J. Eng and Daniel M.T. Fessler. 2007. "Harm, Affect, and the Moral/Conventional Distinction". *Mind and Language* 22 (2): 117-131..
- Kummer, Hans. 1978. "Analogues of morality among non-human primates". In *Morality as a Biological Phenomenon*. Edited by G.S. Stent. University of California Press. 31-47.

- Lillehammer, Halvard. 2003. "Debunking Morality: Evolutionary Naturalism and Moral Error Theory." *Biology and Philosophy* 18 (4): 567-581.
- Nagel, Thomas. 1986. *The View from Nowhere*. Oxford University Press.
- , 1997. *The Last Word*. Oxford University Press.
- , 2012. *Mind and Cosmos: Why the Materialist Neo-Darwinian Conception of Nature is Almost Certainly False*. Oxford University Press.
- Nozick, Robert. 2001. *Invariance: The Structure of the Objective World*. Harvard University Press.
- Olson, Jonas. 2011. "In Defense of Moral Error Theory". In *New Waves in Metaethics*. Edited by Michael Brady. Palgrave Macmillan. 62-83.
- Parfit, Derek. 2011. *On What Matters: Volume II*. Oxford University Press.
- Plantinga, Alvin. 2000. *Warranted Christian Belief*. Oxford University Press.
- , 2011. *Where the Conflict Really Lies: Science, Religion, and Naturalism*. Oxford University Press.
- Railton, Peter. 1986. "Moral Realism". *Philosophical Review* 95 (2): 163-207.
- Rawls, John. 1971. *A Theory of Justice*. Harvard University Press.
- Ruse, Michael. 1998. *Taking Darwin Seriously: A Naturalistic Approach to Philosophy*. Prometheus Books.
- Sayre-McCord, Geoffry. 2009. "Moral Realism". *The Stanford Encyclopedia of Philosophy*. Edited by Edward N. Zalta. URL = <http://plato.stanford.edu/archives/sum2011/entries/moral-realism/>.
- Shafer-Landau, Russ. 2005. *Moral Realism: A Defence*. Oxford University Press.
- Shoemaker, David. 2011. "Psychopathy, Responsibility, and the Moral/Conventional Distinction". *Southern Journal of Philosophy* 49 (S1): 99-124.
- Smetana, J.G. and J.L. Braeges. 1990. "The development of toddlers' moral and conventional judgments". *Merrill-Palmer Quarterly* 36: 329-346.
- Sturgeon, Nicholas. "Harman on Moral Explanations of Natural Facts". *Southern Journal of Philosophy* 24 (S1): 69-78.

- Stevenson, C.L. 1937. "The Emotive Meaning of Ethical Terms". *Mind* 46 (181): 14-31.
- Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127 (1): 109-166.
- , Forthcoming. "Objectivity and Truth: You'd Better Rethink It". *Philosophy & Public Affairs*.
- Timmons, Mark. 2001. *Moral Theory: An Introduction*. Rowman & Littlefield Publishers.
- Toner, Christopher. 2011. "Evolution, Naturalism, and the Worthwhile: A Critique of Richard Joyce's Evolutionary Debunking of Morality". *Metaphilosophy* 42 (4): 520-546.
- Trivers, Robert. 1971. "The Evolution of Reciprocal Altruism". *The Quarterly Review of Biology* 46 (1): 35-57.
- Turiel, Elliot. 1983. *The Development of Social Knowledge: Morality and Convention*. Cambridge University Press.
- de Waal, Frans. 1996. *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*. Harvard University Press.
- , 2000. "Primates – A Natural Heritage of Conflict Resolution". *Science* 28 (289): 586-590.
- Yablo, Stephen. 2001. "Go Figure: A Path Through Fictionalism". *Midwest Studies in Philosophy* 25 (1): 72-102.
- , 2005. "The Myth of the Seven". In *Fictionalism in Metaphysics*. Edited by Mark Eli Kalderon. 88-114.