The Containment Problem and the Evolutionary Debunking of Morality

Tyler Millhouse, Lance S. Bush, & David Moss

Abstract: Machery & Mallon (2010) argue that existing evidence does not support the claim that moral cognition, understood as a specific form of normative cognition, is a product of evolution. Instead, they suggest that the evidence only supports the more modest claim that a general capacity for normative cognition evolved. They argue that if this is the case then the prospects for evolutionary debunking arguments are bleak. A debunking argument which relied on the fact that normative cognition in general evolved might threaten all areas of normative belief, including the epistemic norms upon which the argument relies. For the sake of argument, we accept their claim that specifically moral cognition did not evolve. However, we reject their contention that this critically undermines evolutionary debunking arguments of morality. A number of strategies are available to solve what we call the "containment problem," or how to effectively debunk morality without thereby debunking normative cognition tout court. Furthermore, debunking arguments need not rely on the claim that normative cognition in general evolved. So long as at least some aspects of moral cognition have evolved, this may be sufficient to support an evolutionary debunking argument against many of our moral beliefs. Thus, even if Machery & Mallon are right that specifically moral cognition did not evolve, research in evolutionary psychology may have radical implications for moral philosophy.

1. Introduction

Recent years have seen a dramatic increase in interest in evolutionary debunking arguments (or EDAs').[1] EDAs follow a long tradition of concern about the implications of evolution for morality. Since the rise of evolutionary biology, many have worried that morality may be undermined by the discovery that our moral faculties are, in some sense, a product of evolution. Even Darwin's conviction of purpose in the universe was shaken by "the horrid doubt... whether the convictions of man's mind, which has been developed from the mind of the lower animals, are of any value or at all trustworthy" (Darwin, 1887, p. 316). Contemporary concerns typically do not center on the ultimate purpose of nature (some philosophers have even attempted to exploit biological accounts

[1] For example, Behrends, 2013; Brosnan, 2011; Carruthers & James, 2008; Clarke-Doane, 2012; Cline, 2014; Copp, 2008; De Cruz & De Smedt, 2012; De Lazari-Radek & Singer, 2012; Enoch, 2010; 2011; FitzPatrick, 2013; 2014; Fraser, 2014; Griffiths & Wilkins, 2010; Jong & Visala, 2014; Joyce, 2006; 2013, 2014; Kahane, 2011; Mason, 2010; Peters, 2012; Schafer, 2010; Shafer-Landau, 2012; Skarsaune, 2011; Street, 2006; 2008; Talbott, 2014; Toner, 2011; Vavova, 2014; Visala, 2011; Wielenberg, 2010; Wilkins & Griffiths, 2012.

of function to support ethical theories; e.g., Foot 2001). They do however concern the reliability of human cognition and, more specifically, normative cognition. As we shall outline below, evolutionary debunking arguments can take several forms. Some argue that if our moral beliefs are satisfactorily explained as a product of evolution, then this obviates an account of moral knowledge as grounded in objective, mind-independent truths. Others simply argue that whether there are moral truths or not, if evolution has pervasively influenced our moral beliefs, then we have reason to be skeptical that the moral beliefs we do have are correct.

Machery & Mallon (2010) note that there are at least three ways one might interpret the claim that "morality evolved" which each lend varying degrees of support to evolutionary debunking arguments. While EDAs differ in their particulars, Machery & Mallon argue that they tend to be (at least implicitly) committed to the strongest of these interpretations, namely, that "moral cognition, understood as a special sort of normative cognition," evolved (p. 4). Alternatively, it might be true that "morality evolved" only insofar as a general capacity for "normative cognition- that is, the capacity to grasp norms and to make normative judgments" evolved (p. 4). Finally, it may be the case that "morality evolved" in the sense that merely "some components of moral psychology evolved" (p. 4). Machery & Mallon contend that only the latter two interpretations are supported by the evidence.

If Machery & Mallon are right, then this has profound implications for research into the evolution of morality. It would require both philosophers and scientists to pay closer attention to the metaethical implications of their claims, being careful not to take our present understanding of moral discourse for granted. We do not have the time or resources to fully evaluate the empirical case Machery & Mallon advance in support of their claim. Indeed, Machery & Mallon are careful to present their case as tentative and as dependent on the present state of evidence in an ongoing area of research. As they admit, much important work remains to be done.

Accepting their conclusion for the sake of argument, we evaluate the alleged implications of their thesis for evolutionary debunking arguments. Machery & Mallon hold that if it is not the case that morality evolved in the strong sense that a specifically moral form of cognition evolved, then there is little hope for evolutionary debunking arguments. One of the central problems for EDAs is the possibility that they prove too much. After all, there is no doubt that our capacities to reason about matters far beyond the moral domain have been shaped by natural selection. Are all these capacities thereby debunked? How could one frame an EDA that exempts these capacities? We call this challenge "the containment problem."

Machery & Mallon extend the containment problem by arguing that deprived of the strong claim that "morality evolved," debunkers are left with the more moderate claims: that normative cognition evolved and that components of moral cognition evolved. If debunkers adopt the latter claim, it's not clear that debunkers can demonstrate a pervasive evolutionary influence on morality or that they can obviate truth-tracking accounts. If debunkers adopt the former claim, they risk un-

dermining our broader capacity for normative cognition. This "unpalatable" consequence, Machery & Mallon argue, constitutes an informal reductio ad absurdum of evolutionary debunking.

In the light of Machery & Mallon's challenge, we propose five alternative strategies for constructing evolutionary debunking arguments. If these alternatives are successful, it would no longer be necessary to show that a specific capacity for moral cognition evolved. Rather, research into the details of how evolution has influenced our normative cognition (broadly understood) may suffice to support a strong challenge to many of our moral beliefs. For example, we may be able to ground EDAs in the influence of natural selection on the content of our normative judgments without reference to the moral domain or, indeed, any other normative domain.[2] As a guide to the reader, the structure of this chapter is as follows: In §2, we outline in more detail what is at stake in philosophical discussions concerning how to describe, delineate, and divide the normative domain. In §3, we describe how existing evolutionary debunking arguments attempt, in various ways, to debunk morality. In §4, we outline in greater detail the challenge raised by Machery & Mallon, and the containment problem. This brings us to our more detailed responses to Machery & Mallon's challenge to debunking in §5. Among the responses we propose are: i) 'Revised Domains,' ii) 'Content Domain Debunking,' iii) 'Meta-Normative Debunking,' iv) 'Functional Debunking,', and v) 'Concede and Redeem.' Whether these strategies ultimately succeed is beyond the scope of this chapter. We only intend to evaluate whether they could in principle avoid (or even resolve) the containment problem. Naturally, whether these strategies ultimately succeed depends on further empirical questions about how our cognitive faculties evolved and on the resolution of other philosophical objections. Also, depending on the strength of these alternatives, the success of only one may be sufficient to seriously undermine our moral beliefs as well as Machery & Mallon's reductio. Finally, in §6, we reflect on the implications of these arguments for evolutionary psychologists, suggest areas for future research, and emphasize the need for an Evolutionary Metaethics.

2. Evolutionary Metaethics

A. Meta-normative properties and the moral domain

Machery & Mallon argue that there is no evolved capacity for distinctively moral cognition (p. 20).[3] On their view, moral cognition concerns a particular domain of norms (i.e., moral norms). However, what distinguishes moral norms from other kinds of norms (e.g., social, conventional, religious, and legal norms)? Moral philosophers, psychologists, and neuroscientists have put forward various accounts of the distinctive features of moral norms. Hare (1981) maintains that moral norms differ from other norms in that they provide reasons for action that supersede any competing reasons for action. Bicchieri (2006) proposes that moral/personal norms are unconditional, whereas social

[2]Throughout the paper we equivocate between the influence of evolution and the influence of natural selection. Without offering a full-fledged defense of adaptationism, we assume that the influence of evolutionary mechanisms other than natural selection (e.g., genetic drift) can only enhance EDAs since none of these mechanisms seem suited to producing truth-tracking capacities.

[3]See also, Machery (2012, 2013).

norms are conditioned on the behavior and expectations of others. Gert (2005) claims that moral norms distinctively concern harms to others. Baumard and colleagues argue that moral cognition is fundamentally about fairness (Andr & Baumard, 2011; Baumard et al., 2012; Baumard et al., 2013; Baumard & Sheskin, in press). Still others attempt to identify the moral domain with a distinctive pattern of neural activity (Moll et al., 2005), or a cluster of specialized psychological mechanisms (Haidt, 2012). Finally, Nucci (2001) and Turiel (1983) maintain that humans exhibit a pancultural tendency to reliably distinguish between two distinct classes of norms, one of which corresponds to stereotypical cases of moral norms (e.g. murder, theft) and the other to stereotypical conventional norms (e.g. appropriate dress, table manners).

Which account, if any, is correct is not at issue here. The important point is that different domains of norms are distinguished by their meta-normative properties. These properties simply describe the various ways in which norms can differ from each other. For example, norms can differ in their content, justification, adaptive role, scope of application, authority contingency, seriousness, phenomenology, or social function (Sinnott-Armstrong & Wheatley, 2012). Now that we are clearer on what it would mean to talk about specifically moral cognition, we can ask what it would mean to say that moral cognition evolved. Simply put, moral cognition evolved just in case humans possess an evolved capacity to reason about a certain class of norms (i.e., moral norms).

While we disagree that Machery & Mallon's conclusion seriously undermines debunking arguments, we enthusiastically endorse their efforts to clarify the nature of moral cognition and to distinguish the different ways in which one might interpret the claim that morality evolved. Thus, our goal in this section is to elaborate on Machery & Mallon's argument that moral cognition, interpreted as a distinctive form of normative cognition, did not evolve. We do not intend to endorse this claim, but rather to defend its plausibility. Like Machery & Mallon, we regard as uncontroversial the claim that some of the components involved in moral cognition evolved.

Machery & Mallon's critique of the claim that moral cognition evolved rests on three key points. First, they address the claim that moral norms are universal and that their universality is evidence that they evolved (Hauser, 2006, p. 53; Joyce, 2006, p. 134; Dwyer, 2006, p. 237; Prinz, 2009). Joyce (2006), for instance, claims that the capacity to make moral judgments is present in "all human societies we have ever heard of" and cites the presence of norms in ancient works such as the Epic of Gilgamesh and the Egyptian Book of the Dead as evidence that moral norms seem to appear in every known culture (p. 134). Indeed, even critics of moral nativism such as Jesse Prinz (2009) claim that "Moral norms are found in almost every recorded human society. Like language, religion, and art, morality seems to be a human universal" (p. 167). Yet such assertions run the risk of conflating the uncontroversial claim that every society possesses norms with the much more controversial claim that every society possesses a set of norms that fit the researcher's particular account of what constitutes specifically moral norms (Machery & Mallon, 2010, p. 30). For instance, Prinz (2009) argues that to establish a robust form of moral nativism[4], one approach

[4]Prinz (2009) only considers these the steps necessary to establish a particular form of moral nativism he calls

would be to (1) "identify some moral norms that can be found in all (or almost all) cultures" then (2) "show that innate, domain specific mechanisms are the best explanation of how those norms are acquired" (p. 168). Yet Prinz moves too quickly here. The problem is that this evidence does not show that these cultures possess moral norms in particular. In order to do so, researchers would have to show that the norms found in these cultures share the same meta-normative properties, properties that would allow us to distinguish these norms from non-moral norms; and, as Machery & Mallon are quick to point out, it is unclear that anyone has conducted research that would support such claims.

Second, Machery & Mallon critique one of the most prominent and well-supported lines of evidence that purports to distinguish moral norms from non-moral norms, Turiel and colleagues research on the moral/conventional distinction (Huebner et al., 2010; Nisan, 1987; Nucci, 2001; Nucci & Nucci, 1982; Nucci et al., 1983; Nucci & Turiel, 1978; Smetana & Baeges, 1990; Turiel, 1979; 1983). Naturally, Machery & Mallon could not respond to every account of the moral domain, but they choose a formidable target as an exemplar. According to proponents of the moral/conventional distinction such as Nucci (2001) and Turiel (1983), human beings exhibit a pancultural tendency to reliably distinguish between moral and conventional norms. Machery and Mallon (2010, p. 32) note that moral norms are supposedly judged to be independent from authorities, to hold universally, and to be grounded in harm to others, infringement of rights, or justice, whereas conventional norms depend on authority, are locally applicable, and are grounded in local practice. There are two questions raised by this model: (1) is the distinction innate, and (2) is the distinction actually reflected in people's judgments?

Much of the work in favor of the moral/conventional distinction is developmental. The reliable and early emergence of the distinction has prompted some to conclude that it is innate (e.g. Dwyer, 1999, 2006; Joyce, 2006; Mikhail, 2000; Wilson, 1993). These arguments, like many nativist arguments, typically appeal to a poverty of the stimulus'. Interestingly, Nucci (2001) explicitly rejects nativist accounts, arguing that children learn the moral/conventional distinction by attending to whether the consequences of actions are fixed (e.g., hitting and pain) or whether they are socially/culturally mediated (e.g., vulgar language and offense).[5] Machery & Mallon cite additional work which questions whether children actually do reliably draw the distinction (e.g., Gabennesch, 1990; Carter & Patterson, 1982).

Machery & Mallon further criticize the distinction by citing work that calls into question whether adult judgment conforms to the moral/conventional distinction. For example, Kelly et al. (2007) report that judgments about several harm norms appear to display authority dependence and cultural relativity. In other cases, individuals will judge an act to be seriously wrong and authority independent, but fail to justify their conclusion by appeal to harm, rights, or justice (Haidt et al.,

"immodest moral nativism" (p. 168). Thus, Prinz acknowledges that there are other means by which one could establish some form of moral nativism.

[5]Turiel (1983) also disputes whether the developmental evidence supports nativism.

1993). There are certainly reasons to doubt these studies (e.g., Rosas, 2012), but combined they do suggest that the moral domain may be less distinct and less unified than previously supposed.

Third, Machery & Mallon criticize other developmental evidence that has been cited in support of the claim that specifically moral cognition evolved. For example, they note that it is tempting to conclude from the evidence that young children seem able to understand deontic conditionals (e.g., "It's not safe outside for the squeaky mouse, so all squeaky mice must stay in the house") earlier and more easily than indicative conditionals (e.g., "It's not safe outside for the squeaky mouse, so all squeaky mice are in the house") that moral cognition specifically evolved (2010, p. 52). Yet they argue that this simply suggests that a special capacity for normative cognition appears early and has evolved, not that a tendency for moral cognition specifically evolved (since the deontological conditionals in question are not distinctively moral, merely normative) (2010, pp. 52-53).

Similarly, they criticize the inference made by Dwyer (2007) from the fact that infants seem to develop a tendency for empathetic responses to others' suffering from an early age, to the conclusion that this is compelling evidence that moral capacities develop at an early age. Again, they note that empathetic responses to suffering need not be taken to be distinctively moral, even if empathy is often thought to be morally significant (Machery & Mallon, 2010, p. 53). Such capacities (for empathetic responses) might be evolved psychological traits, and might be recruited by moral cognition, but it does not follow that moral cognition specifically evolved.

B. Metaethical variability and indeterminacy

A second line of evidence that provides indirect support for Machery & Mallon's rejection of the claim that distinctively moral cognition evolved comes from a growing body of research on the psychology of folk metaethics. In particular, experimental metaethics suggests that folk intuitions about the meta-normative properties of norms exhibit both interpersonal and intrapersonal variability. Insofar as such evidence suggests that putatively moral norms do not share a distinctive set of meta-normative properties, this evidence lends some indirect support to Machery & Mallon's claim that specifically moral cognition did not evolve. The reason for this is that, if there is no distinct moral domain, then this tends to undermine the notion that there is a dedicated, innate faculty for reasoning about moral norms.

Of course, the question of the unity of the moral domain is strictly independent of the question of its distinctiveness. For example, despite variations in people's meta-normative judgments about moral norms, there may nevertheless be relatively sharp distinctions between moral norms and other norms. For example, moral norms might all possess at least three of four meta-normative properties. This would imply that there are at most five types of moral norms (i.e., 1,2,3,4; 2,3,4; 1,3,4; 1,2,4; 1,2,3). In this case, there would be no property that all moral norms share. However, we could still distinguish these norms from norms that have less than three of these properties. That

said, disabusing ourselves of the notion that all moral norms share a distinctive set of properties should make us less convinced that the boundaries between moral norms and other kinds of norms are adequately sharp. If even some meta-normative properties are shared between different kinds of norms, then we must ask whether or not reasoning about those norms is subserved by the same cognitive mechanisms.

Gill (2009) refers to the assumption that moral discourse can be captured by a single, unified account as the Uniformity-Determinacy (or 'UD') assumption. According to Gill, the UD assumption may be false. Rather than being uniform and determinate, ordinary moral discourse could be variable and indeterminate. Gill refers to this as the Metaethical Indeterminacy-Variability (or 'IV') thesis. Variation may occur within or between individuals. Intrapersonal variability occurs when the same individual makes moral judgments in a way that fits one account in some cases but a different account in others. For example, the same individual might presuppose one set of meta-normative properties when reasoning about some moral issues but a different set when reasoning about other moral issues (e.g., by assuming relativism when thinking about sexuality, but assuming absolutism when thinking about physical harm.) Interpersonal variability or intergroup variability occurs when different individuals or groups presuppose different meta-normative properties when making moral judgments. For example, one cultural group might understand morality as authority independent, whereas another might understand morality to derive from divine commands.

The IV thesis derives much of its plausibility from anecdotal observations that people report an adherence to conflicting metaethical standards or speak and act in ways that suggest that they do. However, Nichols (2004) has also found empirical support for such observations. Across five studies, most of Nichols's subjects treated moral claims as objective, yet a third or more gave responses that suggested that they were moral relativists. Subjects were also asked to explain their answers, and often gave explicitly relativist rationales. For example, one subject stated that "Morality is subjective to culture," but that whether the earth is flat is "a cold, unwavering scientific fact" (as quoted in Nichols, 2004, p. 4). Other studies provide evidence of context variability and content variability (Beebe, 2014; Goodwin & Darley, 2008; 2012; Sarkissian et al., 2011; Wright et al., 2013). Sarkissian et al. (2011) found that the greater the cultural difference between two hypothetical people who disagreed about a moral issue, the more subjects agreed that both of these people could be right. As Beebe (2014) observes, these findings suggest that "folk intuitions about metaethical objectivity vary as a function of cultural distance, with increased cultural distance between disagreeing parties leading to decreased attributions of metaethical objectivity" (p. 167). In addition, researchers have found that the perceived objectivity of moral norms is in part a function of their content, with some norms such as robbing a bank exhibiting a high degree of perceived objectivity, whereas others, such as abortion and assisted suicide, are perceived as less objective (Goodwin & Darley, 2008; 2012). These findings are far from decisive evidence of the variability thesis, but they at least shift some of the burden of proof back on proponents of the UD assumption (Gill, 2009). Once again, it is possible that, however disunified, moral norms

are distinct from other kinds of norms. Nevertheless, as morality is revealed to be less and less unified, it seems increasingly likely that morality will be shown to share important features and even cognitive mechanisms with other kinds of norms.

3. Varieties of Debunking Arguments

As stated before, our primary project in this chapter is not to fully evaluate the conclusion of Machery & Mallon (2010), but to evaluate what their conclusion implies for evolutionary debunking arguments (and ultimately for evolutionary psychology). Evolutionary debunking arguments attempt to use the evolutionary origins of moral cognition to undermine moral realism. Moral realism is the view that moral claims attempt to report mind-independent facts (as opposed to, say, emotions or attitudes), and that at least some of these claims are true (Sayre-McCord, 2009).[6] There are several ways in which evolutionary theory might undermine realism, and some are more controversial than others. Typically, debunkers do not attempt to demonstrate that there are no objective moral truths–demonstrating non-existence is notoriously difficult. Rather, debunkers typically attempt to epistemically undermine moral realism, that is, they attempt to undermine our evidence in favor of the realist position.

The most basic form of debunking involves identifying and checking the empirical claims of moral philosophers. Just which of a philosopher's claims should be understood as empirical is not always obvious, but once these claims have been identified, there is generally no question as to whether one may raise a scientific challenge. Insofar as certain empirical claims are central to an ethical theory and concern evolution, evolutionary science may support debunking arguments against that theory. Less directly, some varieties of moral realism (e.g., Foot, 2001) make use of notions ostensibly drawn from evolutionary biology (e.g., 'function', 'species', or 'fitness'). Insofar as these theories depend on the currency of these notions in evolutionary science, we can assess how faithfully ethicists have interpreted the science and whether the concepts on which they rely actually enjoy the support of the best scientific theories. If the conceptual foundation of an ethical theory is sufficiently undermined, that theory is debunked.[7]

By contrast, most realist theories make no essential reference to evolutionary theory, and the most important (and controversial) debunking arguments go beyond mere fact-checking. Unless

[6] By mind-independent we don't mean to say that moral facts cannot mention mental states. For example, causing unnecessary pain might be morally wrong, but it is not wrong in the realist sense if its truth depends on how people regard that prohibition. For example, whether evolution occurred does not depend on how people regard the theory. On the other hand, whether something is money depends crucially on whether people regard that thing as a medium of exchange.

[7] An illustrative analogy is the pseudo-scientific use of scientific concepts. When new age healers talk about quantum uncertainty or the vibration of strings, the healers intend to support their theories by illustrating their coherence with or basis in established science. Closer inspection reveals (of course) that the notions employed by healers bear little resemblance to their scientific counterparts. Once one has substituted the genuine notions for the fakes, the theories of pseudo-scientists generally lose whatever facial plausibility they once enjoyed. Of course, this is an extreme example, but similar problems may exist in philosophy. Ladyman, et al. (2007) levy a similar critique against the dubious use of physical concepts by metaphysicians (p. 25-27).

otherwise noted, all references to "evolutionary debunking arguments" in this chapter refer to these kinds of arguments. Within evolutionary debunking arguments proper, there are roughly two approaches, arguments from influence and arguments from sufficient non-moral explanation.

A. Influence

Influence-based arguments contend that evolution has pervasively influenced the content of our moral judgments[8] by processes unconcerned with moral truth. This argument is meant to undermine our justification for accepting the conclusions of our moral judgments. By way of analogy, suppose we learned that a private interest in maximizing publications motivates many researchers, this should undermine our confidence in the content of academic publications, ceteris paribus. We trust these sources because there are constraints that help make epistemically virtuous research practices an instrumental goal of the selfish researcher (e.g., peer-review and severe consequences for plagiarism and fabrication). The problem in this case is not that a desire to publish is strictly opposed to arriving at true findings. Rather, the problem is that this desire is indifferent to true findings, except insofar as such findings further the goal of publication. By a similar argument, debunkers hold that if our capacity for moral judgment has been pervasively shaped by natural selection, this should undermine our trust in the content of our moral beliefs, ceteris paribus. Natural selection is unconcerned with producing true beliefs except insofar as true beliefs enhanced ancestral fitness. Of course, if one could demonstrate either (1) that we have no reason to believe that evolution has significantly shaped our moral capacities, or (2) that there are reasonable grounds for expecting our evolved moral capacities to track the truth instrumentally, then this brand of debunking can be defeated.

B. Sufficient Non-Moral Explanations

Sufficient non-moral explanation arguments contend that truth-tracking theories of moral judgment are superfluous. One influential philosophical theory (Quine, 1948) holds that we should be committed to the existence of just those kinds of things that the best scientific theories tell us exist. Regardless of how strictly one adheres to this theory of ontological commitment, it's hard to deny that a simpler, more explanatory theory should supplant a more complicated, less explanatory one. Importantly, we generally have no problem concluding (if tentatively and probabilistically) that belief in the entities and processes proposed by superseded theories is unwarranted.

Some varieties of moral realism hold that moral facts are true in virtue of natural ones (ethical naturalism). Others hold that moral facts are true in virtue of non-natural ones (ethical non-naturalism). In either case, we can imagine different kinds of explanations for our moral

[8] In general, when we speak about moral judgments we intend to include what philosophers' call "moral intuitions." In ordinary language, intuition' often refers to a gut decision or feeling. While moral intuitions may frequently take this form, philosophers typically see intuitions as including judgments that involve more explicit reasoning or consideration.

judgments–some that explain moral judgment as tracking (or attempting to track) a domain of moral facts and some that explain moral judgment without reference to a domain of moral facts. If the best theories turn out to be of the latter sort, there seem to be good grounds for dropping our commitment (if we had one) to a domain of moral facts, since the notion of such a domain belongs to a superseded explanation. Insofar as evolutionary theory provides a superior non-truth-tracking theory, it can offer grounds for rejecting moral realism.

There are several objections one might levy against these approaches–either in their schematic form or as applied in particular arguments. Many of these objections are beyond the scope of our present discussion. However, a new objection to debunking arguments suggests serious implications, not only for debunking arguments but for broader work on the evolution of morality.

4. The Containment Problem: Debunking Debunked?

A. The Containment Problem

A popular argumentative strategy across disciplines is reductio ad absurdum. Formally, a reductio is a proof that a proposition is false because it entails a contradiction. Informally, a reductio is an argument to the effect that a proposition is false because it entails an absurd or unlikely conclusion. For example, if an ethical principle requires one to dine nightly on human infants, one would (quite reasonably) count that as a strong reason to reject the principle in question. The primary benefit of reductios is that they are indirect. Rather than directly rebutting the reasons one's opponent has offered in favor of some proposition, p, one can cast doubt on those reasons by showing that if they support p, they also support some further, untenable conclusion. Unlike formal reductios which involve strict contradictions, informal reductios always allow one to accept the unsavory implications in the name of preserving one's original conclusion.[9]

One such argument against evolutionary debunking attempts to show that the central premise in debunking arguments has undesirable implications beyond morality. Plantinga, for rather different purposes, concisely summarizes this point:

> Evolution is directly interested (so to speak) only in adaptive behavior (in a broad sense including physical functioning), not in true belief. Natural selection doesn't care what you believe; it is interested only in how you behave. It selects for certain kinds of behavior: those that enhance tness, which is a measure of the chances that one's genes will be widely represented in the next and subsequent generations. It doesn't select for belief, except insofar as the latter is appropriately related to behavior. But then the fact that we have evolved guarantees at most that we behave in certain ways... The

[9]Kant (in)famously argues that one should not lie to a murderer to protect the innocent. This is an unsettling consequence of his deontological ethics, but one he appears willing to accept. Others have not been as willing to follow Kant in accepting this implication, with some insisting that Kant's own principles do not require that one always tell the truth (Korsgaard, 1986) and others accepting that it does (Constant, 1796) and concluding that this is a deeply problematic implication for Kant's ethical system.

> objective probability that our cognitive faculties are reliable, given naturalism and given that we have been cobbled together by the processes to which contemporary theory calls our attention, is low. (2009, p. 302)

The important point here is that the evolutionary debunker must provide some grounds for thinking that the evolutionary influence on moral judgments is importantly different (in either character or strength) from the evolutionary influence on other judgments which the debunker considers reliable. The challenge of drawing these distinctions we call 'the containment problem'.

One traditional solution to the problem is to suggest that our capacity for moral judgment contributed to our ancestral fitness in a different way than, say, our capacity for perceptual judgments. For example, we have good reasons to expect biases in perception which favor the least costly error (as assessed in the EEA)(Haselton & Nettle, 2006). Why aren't these biases more pervasive? What if they are and we cannot tell? We cannot (of course) give a complete response to these worries, but we can identify some relevant constraints. First, there are costs associated with responding to false alarms (e.g. in time and energy). Second, there are cognitive costs arising from biases. A bias towards detecting certain objects necessarily makes other perceptions less likely. For example, if one saw a bear in every cave or a tiger in every shadow, one couldn't identify cases where danger was absent even when doing so would have been ancestrally adaptive. While these constraints still permit substantial bias, they make it harder to imagine the evolution of a pervasively misleading perceptual system.[10] While plausible truth-tracking theories are easy to generate for perception, they are harder to generate for moral judgment. Whether or not these moves are ultimately successful for debunkers, the key point here is that the they are attempting to resolve the containment problem by reference to relevant differences between moral cognition and our other capacities.

B. The Metaethical Containment Problem

As detailed in §2, Machery & Mallon (2010) argue that the claim 'morality evolved,' when interpreted in its strongest form (as the claim that specifically moral cognition evolved), overreaches the available evidence. As a substitute for this claim, they contend that the less specific claim 'normative cognition evolved' is more defensible. The idea here is that however we come to make moral judgments, our competence is not the result of an evolved system for moral cognition specifically. Rather, moral reasoning is the result of a general system for normative cognition that can be exploited to subserve reasoning about a contingent and culturally-variable moral domain.

One of the philosophical implications suggested by Machery & Mallon is for evolutionary debunking arguments (detailed above). They contend that insofar as debunking relies on the claim that "morality evolved" (i.e., moral cognition evolved), it is undermined by their conclusion. It's

[10] Obviously more needs to be said here, but a lengthy digression into the evolution of perception seems inappropriate.

hard to see at first glance how this is so. After all, the claim that normative cognition evolved appears to be broader than the claim that morality cognition evolved, since moral cognition is a type of normative cognition. Thus, the conclusion that normative cognition in general evolved seems to license (if anything) a broader critique of normative reasoning. However, this is precisely the problem. If evolution undermines moral realism, Machery & Mallon argue, then it must do so via the evolution of this broadly normative system (since that's all that really evolved). This makes it hard to see how other normative domains can survive unscathed. They contend:

> We have focused on a second interpretation of the claim that morality evolved: normative cognition—the capacity to grasp and apply norms—evolved...this conclusion is cold comfort to those philosophers who want to get some philosophical mileage out of evolutionary findings. This is particularly clear when one focuses on the argument that the evolution of morality would undermine the authority of moral norms [...] Suppose that this argument from the evolution of morality is meant to hang on the reading of the claim that morality evolved considered in this section: normative cognition in general evolved... If the evolution of normative cognition really undermines the authority of moral norms, then it should also undermine the authority of any kind of norms (including epistemic norms), for there is no reason why only the authority of moral norms would be undermined by the evolution of the capacity to grasp norms tout court. (p. 19)

If Machery & Mallon are correct, then the strategy proposed in the previous section fails since it depends on the distinctive evolutionary history of moral cognition. Of course, they do assume that debunkers care about defending realism in other normative domains. In particular, Machery & Mallon cite epistemic norms as a point of special concern. To undermine those, it seems, would undermine any grounds one might have for endorsing a debunking argument in the first place. Going forward, then, we will generally assume (1) that distinctively moral cognition did not evolve and (2) that some of our normative judgments (esp. epistemic norms) are worth defending.

5. Alternative Debunking Strategies

Put briefly, our conclusion is that Machery & Mallon's evolutionary claim is merely cool comfort to the evolutionary debunker. It certainly rules out nave approaches to debunking which take for granted a determinate, invariable, and innate target for debunking (i.e., moral cognition). However, we believe that a number of powerful strategies remain open to the debunker. Our purpose here is not to pass judgment on the ultimate success of these strategies against the realist, but rather to evaluate them in light of the metaethical containment problem.

i. Revised Domains

The debunker needn't defend evolved, containment-supporting distinctions between every

domain of normative cognition. Suppose, for instance, that the moral/conventional distinction is not innate/evolved but varies with culture, education, or class. This needn't trouble the debunker unless he or she is concerned with defending realism in the conventional domain. The only distinctions the debunker needs to defend are those between the targeted domains and the defended domains. For example, we might imagine a distinction between behavioral norms and norms of rationality. Behavioral norms' might encompass what we commonly consider moral, social, and conventional norms, with "norms of rationality" encompassing epistemic norms and the norms of instrumental reasoning. If the debunkers can construct a successful debunking argument against moral realism and aren't concerned with defending realism in the social and conventional domains, they should not be troubled if their argument debunks the entire domain of behavioral norms.

This kind distinction would hardly be unprecedented in the philosophical literature. On Hume's sentimentalist ethical theory, reason alone is not a source of motivation. Reason, for Hume, is a "slave of the passions" and while it can determine how to achieve what one desires, it cannot determine what one desires (Hume, 2007; Cohon, 2010). While both norms of rationality and behavioral norms prescribe behaviors (in a sense), it is only behavioral norms that non-instrumentally guide motivation. For example, the norm update your beliefs according to Bayes' theorem is neutral with respect to an agent's utility function, but the norm do not murder is not. It is interesting that many criteria for rationality presuppose a system of preferences or desires on the part of an agent (Russell, 1997). Much more needs to be said about this to fully flesh out the behavioral/rational distinction, but it does serve as a useful example of one possible approach to revising our normative domains. Ultimately, further empirical work is needed to determine whether Hume's conjecture about the motivational roles of sentiment and reason is psychologically plausible.

ii. Content Domain Debunking

The claim that moral reasoning per se did not evolve is consistent with the claim that the content of our normative beliefs has been influenced by evolution. For example, suppose we found that some cultural groups identified the prohibition of incest as a conventional norm and others identified it as a moral norm. In addition, suppose that having some kind of norm against incest is nearly universal. One interpretation of this finding is that while we have an evolved, innate bias towards condemning incest, we have no innate bias concerning the normative domain under which we condemn it. Indeed, we might imagine whole content domains (fairness seems a likely candidate) that exhibit pervasive evolutionary influence. These content domains might primarily encompass normative judgments we consider moral, but also encompass normative judgments we regard as conventional, personal, social, or legal. Perhaps these content domains could serve as a target for debunking.

Private property intuitions provide a felicitous example. Property rights appear to straddle several normative domains (at least as we conceive them). Issues of property rights appear in moral, social, and legal norms. These norms emerge fairly early in human development (e.g., in preschool

children) (Bakeman & Brownlee, 1982; Weigel, 1984). In addition, as Gintis (2007) observes, property norms are also present, mutatis mutandis, in a wide range of nonhuman species, from butterflies to apes. These norms often employ some standard of prior possession or occupation.

Gintis (2007) and many others (e.g., Dawkins, 2006; Maynard Smith & Parker, 1976) suggest that evolution exploited these salient local features (e.g., occupation/possession) to support a novel strategy for handling resource conflicts. To see why this is the case, we can begin by considering only aggressive (i.e. "hawk") and submissive (i.e., "dove") strategies in a resource conflict game. The result is a mixed strategy equilibrium. Once individuals can recognize occupation/ possession, it can then be used as a signal in a correlated equilibrium (in this case a "property equilibrium", Gintis, 2007, p. 3), and deference to possessors/occupiers can invade the population. Strictly speaking, deference to invaders could also be stable, but this is relatively rare in actual populations (Dawkins, 2006). Given these considerations, the observational and theoretical case for an evolved basis for our property intuitions appears quite strong. Of course, nothing in the account above invokes moral truth. All this plausible history does is estimate the payoffs of different strategies in terms of individual reproductive fitness and use these payoffs to explain population-level phenomena (i.e., the observed frequencies of different strategies).

Let's stipulate that this selective history of property rights judgments undermines or obviates a truth-tracking account of these intuitions (i.e., it debunks them[11]). Is this a moral debunking argument? Not as such, since it doesn't target norms according to their particular normative domain (e.g., the moral domain). As we noted earlier, our intuitions about property seem to shape norms across domains. It would be better to see this as content domain debunking. As such, it evades the problems raised by Machery & Mallon for normative domain debunking. By using our normative intuitions about property as an example, we hope to show just how ambitious content-based debunking could be. As it happens, debunking arguments which selectively target other content domains have already been offered. Lazari-Radek & Singer (2012) argue that one can selectively debunk partialist ethical intuitions. Partialism is the view that (in principle) it is morally justified or even obligatory to prefer those one knows or with whom one has a particular social or genetic relationship. Impartialism does not deny that we may have instrumental reasons to act in the interest of those with whom we have greater familiarity. After all, we might often be in a better position to aid those whose needs we know best. However, impartialism denies that we have any non-instrumental (i.e. principled) reasons to prefer those with whom we happen to be acquainted. In any case, it seems that most people have strong partialist intuitions (esp. in regard to parental obligations). Lazari-Radek & Singer argue that normative intuitions in favor of partialism are quite plausibly a product of natural selection (e.g., via kin selection), whereas impartialist intuitions defy evolutionary explanation. Hence, only partialist intuitions are subject to evolutionary debunking. There are reasons to believe that evolution can account for key aspects of our impartialist judgments

[11]It's important to note that this debunking account would only undermine those property-related normative judgments that depend (directly or indirectly) on the relevance of possession or occupation to ownership (e.g., approving of adverse possession as a means of acquiring property).

(Kahane, 2014), but the debunking of partialism can proceed regardless.

iii. Meta-Normative Debunking

As discussed earlier, some researchers have attempted to individuate normative domains (in part) by identifying regularities in the attribution of different meta-normative properties to different subsets of norms. Suppose (following Machery & Mallon) that these projects will fail to identify a universal moral domain. Nevertheless, there may still be a universal, evolved capacity to make certain kinds of meta-normative judgments.

For example, suppose there were plausible reasons to believe we have a specially evolved capacity to reason about authority independent norms. We might be able to investigate these reasons without invoking a normative domain within which independence judgments tend to occur. Indeed, the common factor eliciting such judgments may not be anything proper to the norm in question. For example, perhaps such judgments arise in circumstances of conflict with social superiors (regardless of the nature of the particular normative judgment under dispute). The capacity to invoke an authority independent answer to normative questions may confer a social or rhetorical advantage. This is purely speculative, of course, but the point is that our capacities to reason about authority contingency, generalizability, etc. may have selective histories in their own right. Depending on the details, these histories might support debunking. For example, if one could substantiate the claim that our capacity for reasoning about authority independence evolved and cannot identify any plausible truth-tracking account of authority independence judgments, then perhaps such judgments should not be trusted.

iv. Functional Debunking

There is no reason that the distinctions found in folk morality need to track the joints between underlying cognitive mechanisms. For example, if the dual process theory of moral judgment is correct, what many regarded as "moral judgment" encompasses two cognitive processes (Greene, 2007; Cushman & Greene, 2012; Cushman et al., 2010) which tend to produce different kinds of moral judgments. Similarly, we might discover (as seems plausible) that many of the mechanisms underlying epistemic judgment are functionally distinct (in at least some important ways) from the mechanisms underlying other forms of normative cognition. If this turns out to be the case, then regardless of whether epistemic norms are distinct from moral norms in folk metaethics, there are plausible grounds for thinking about them differently with respect to their reliability since they depend on distinct cognitive mechanisms. Also, a functional distinction between epistemic and moral judgment processes would suggest separate evolutionary histories for each which may be debunking (or not) depending on their individual elements.

Of course, if only some of the mechanisms that contribute to a moral judgment are evolved, this may be insufficient to support debunking (as Machery & Mallon suggest). This kind of debunk-

ing would depend on a class of normative judgments being substantially influenced by a particular cognitive mechanism. For example, on Greene's view, automatic/affective processing determines moral judgments that differ from moral judgments made under cool(er) controlled processing. For Greene, each of these dual-processes depends to some degree on affect. Only the goals of controlled processing are set via affect (similar to the Humean picture), but in the case of automatic processing, the immediate practical content of the judgment (i.e., what you ought to do) is determined by affect. If the system that governs our affective responses to moral situations is innate and evolved, one might be able to run a debunking argument against the output of automatic processing, but not against the output of controlled processing.

v. Concede and Redeem

The debunker might simply accept that no normative domain is spared debunking. This is less radical than it first appears. Let's suppose that the evolution of normative cognition does provide some reason to doubt the conclusions of our normative reasoning tout court. That's far from the only data we have. There might be other good reasons to think that aspects of our normative reasoning are sound. After all, it is not that a process of evolution cannot produce a truth-tracking system, it is merely that it is strictly indifferent to doing so. For example, we can ask with Wigner why mathematics is so "unreasonably" effective (1960). We can ask why science has been such a remarkable success (Putnam, 1975). Our best explanation in such cases may be that these disciplines (with the aid of the normative judgments they require) are tracking something deep about the world. Of course, in adopting this approach, the debunker must commit to redeeming whatever parts of normative reasoning he or she intends to protect from debunking. These redemption arguments can follow the same strategies outlined for debunking. For example, the debunker might be able to selectively redeem our reasoning in a particular content domain, in a stipulated normative domain, or in a revised (but natural) normative domain. The debunker might also selectively redeem normative reasoning that depends on reliable cognitive mechanisms. Of course, the realist might offer similar defenses of those aspects of normative reasoning targeted by the debunker. Whether any of these defenses are ultimately successful is a separate matter–we merely intend to point out another path open to debunkers.

6. Some Implications for Evolutionary Psychology

A. Talking About the Evolution of Morality

So far, we have disputed the philosophical implications of a hypothesis in evolutionary psychology. In addressing those implications, we have outlined strategies for philosophers interested in evolutionary debunking. However, Machery & Mallon's hypothesis has wide ranging implications for evolutionary psychologists working on the evolution of morality. We believe that many of the strategies outlined for the debunker can be adapted for use in evolutionary psychology. The reason

for this is straightforward. Both debunkers and evolutionary psychologists are interested in characterizing the influence of evolution on our normative judgments. While we highlight the relevance of functional debunking' and content domain debunking', other strategies suggest similar ways of characterizing the evolutionary origins of normative cognition. For example, one might talk about the evolution of our capacity for specific meta-normative judgments (suggested by meta-normative debunking'), or one might propose new normative domains that more plausible share a common evolutionary origin (suggested by revised domains').

i. Functional Debunking

The third strategy is perhaps the most relevant to researchers in evolutionary psychology. In fact, modular approaches to cognitive processes are especially suited to this strategy. Essentially, instead of talking about the output of innate mechanisms as moral judgments, one might refer to these judgments as simply normative, allowing the modules to individuate their outputs. For example, a researcher working within Greene's (2007) dual process model of moral cognition might distinguish between judgments that arise from each process and talk separately about each.

This approach also invites greater specificity about the contribution of different cognitive processes to our meta-normative judgments. For example, on Nichols' (2004) sentimental rules account, representations of rules are distinct from relevant affective responses. Certain affective responses to rule violations (e.g., anger about harm) mediate judgments about the generality of the rule (i.e., how widely it applies). For example, he hypothesizes that strong negative affect disposes individuals to more broadly generalize about their judgments (e.g., that action is always and everywhere wrong). Nichols' model explains why strong affect (e.g., disgust at the sight of a surgery) does not automatically trigger condemnation and why condemnation is not always severe or absolute. In any case, the functional division between rule representation and affective processing would allow researchers to study rule representation without the metaethical baggage of invoking specifically moral rules. Hence, when meta-normative judgments are at issue, researchers could more precisely indicate which meta-normative judgments are produced by which systems if they drop references to morality.

ii. Content Domain Debunking

Nativism about moral cognition can be separated from nativism about the content of normative judgments. In other words, it is possible to argue for biases in normative judgment that concern which norms are endorsed, but not how those norm are endorsed (e.g., as authority independent, or culturally relative, etc.). There might be any number of content biases in ordinary judgment. As discussed in the context of content domain debunking, one might argue that people are innately predisposed to endorse norms that require care for family. These norms might be viewed as social norms in one society and moral norms in another. They might even be represented as a kind of

norm that doesn't fit into any of the normative domains we've so far discussed. However, in nearly all cultures, there may be norms (of one kind or another) that require caring for one's family. The best explanation for this universality may be a nativist one. Such an explanation need not invoke morality in particular, only a bias in broader normative cognition.

B. Areas for Future Research

All of the suggestions we've made for the debunker depend crucially on the outcome of relevant empirical research. While there are some extant evolutionary accounts that support these approaches, there are some particular areas for future research that would be relevant to debunkers (and those who wish to question their empirical presuppositions). The most obvious area for future research is on the particular questions raised and conclusions drawn by Machery & Mallon. However, we also attempt to highlight a more specific area for future research.

i. The Evolution of Meta-Normative Reasoning

Even if Machery & Mallon are correct that there is no set of meta-normative properties that universally typifies the moral domain, there remain important questions about how early humans developed the capacity to attribute these properties to norms and normative domains. Consider the notion of categorical ought claims (contrasted with a hypothetical ought claims):

> Hypothetical ought claim: If do not want to feel guilty, you ought not steal.
> Categorical ought claim: You ought not steal.

It might be easy to take this difference for granted, and imagine the acquisition of the categorical notion by simply dropping the antecedent. However, it's not clear either that people would be inclined to do this or that the result would be comprehensible. Note, we don't consider it acceptable or meaningful to say If you do not want to feel guilty', why should the reverse be any different? Indeed, some philosophers suggest that it is not (Neurath, 1983, p. 54). In any case, the why and how of our facility with categorical ought claims presents interesting questions for researchers.

To give a more concrete example, let's continue with example given earlier– Nichol's suggestion that strong negative affect mediates the generalization of moral norms. The idea here is that strong negative affect might incline people to universalize their normative judgments. Perhaps there are features of circumstances that give rise to negative affect that also made a tendency to generalize adaptive. For example, if endorsing a norm as true for everyone enhances other's tendency to comply with that norm, then endorsing and defending a general reading of certain norms may be advantageous. For example, it might not be wise to conditionalize murder norms[12], since a higher risk of being murdered would not have been ancestrally adaptive ceteris paribus. Applied

[12]Exactly what murder means would have to be fleshed out. It would be crucial to see which forms of killing are permitted and under what circumstances.

more broadly, the idea might be that the tendency to generalize tracks actions (via negative affect) that were especially serious risks in the EEA.

Again, the main idea here is that evolutionary researchers can shine light on our narrower metaethical competences. This project avoids questions about the unity and universality of the moral domain, by instead looking at meta-normative properties apart from the domains they traditionally characterize. Naturally, this will be a difficult project and conclusive evolutionary accounts are unlikely, but that doesn't preclude significant advances in our understanding.

7. Conclusion

In this chapter, we have presented Machery & Mallon's (2010) argument in favor of the claim that morality, as such, did not evolve. In addition, we have explored the implications of this conclusion for evolutionary debunking arguments and offered alternative debunking strategies that do not rely on the contested claim. In so doing, we have attempted to identify new ways of talking about the evolution of morality that both avoid controversial metaethical commitments and encourage a more precise and explicit discussion of our capacity for normative reasoning. Ultimately, we hope to encourage additional work in empirical evolutionary metaethics that illuminates the origins and structure of normative cognition.

References

Andre, J. B., & Baumard, N. (2011). Social opportunities and the evolution of fairness. *Journal of theoretical biology*, 289, 128-135.

Baumard, N., & Sheskin, M. (2015). Partner choice and the evolution of a contractualist morality. In J. Decety & T. Wheatley (Eds.), *The moral brain*. Cambridge, MA: MIT Press.

Bakeman, R., & Brownlee, J. R. (1982). Social rules governing object conflicts in toddlers and preschoolers. In *Peer relationships and social skills in childhood* (pp. 99-111). New York, Springer.

Baumard, N., Andr, J. B., & Sperber, D. (2013). A mutualistic approach to morality: The evolution of fairness by partner choice. *Behavioral and brain Sciences*, 36(01), 59-78.

Baumard, N., Mascaro, O., & Chevallier, C. (2012). Preschoolers are able to take merit into account when distributing goods. *Developmental psychology*, 48(2), 492.

Beebe, J. (2014). How different kinds of disagreement impact folk metaethical judgments. In J.C.

Wright & H. Sarkissian (2014), *Advances in experimental moral psychology: Affect, character, and commitment* (pp. 167-187) London: Continuum.

Behrends, J. (2013). Metanormative realism, evolution, and our reasons to survive. *Pacific philosophical quarterly*, 94(4), 486-502.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms.* Cambridge: Cambridge University Press.

Brosnan, K. (2011). Do the evolutionary origins of our moral beliefs undermine moral knowledge?. *Biology & philosophy*, 26(1), 51-64.

Carter, D. B., & Patterson, C. J. (1982). Sex roles as social conventions: The development of children's conceptions of sex-role stereotypes. *Developmental psychology*, 18(6), 812.

Carruthers, P., & James, S. M. (2008). Evolution and the possibility of moral realism. *Philosophy and phenomenological research*, 77(1), 237-244.

Clarke-Doane, J. (2012). Morality and mathematics: The evolutionary challenge. *Ethics*, 122(2), 313-340.

Cline, B. (2014). Nativism and the evolutionary debunking of morality. *Review of philosophy and psychology*, 1-23.

Cohon (2010). Hume's moral philosophy. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 ed.). Retrieved from http://plato.stanford.edu/entries/moral-realism/

Copp, D. (2008). Darwinian skepticism about moral realism. *Philosophical issues*, 18(1), 186-206.

Constant, B. (1996). Des réactions politiques. Retrieved from http://classiques.uqac.ca/classiques/constant_benjamin/des_reactions_politiques/des_reactions.html

Cushman, F., & Greene, J. D. (2012). Finding faults: How moral dilemmas illuminate cognitive structure. *Social neuroscience*, 7(3), 269-279.

Cushman, F., Young, L., & Greene, J. D. (2010). Our multi-system moral psychology: Towards a consensus view. in J. Doris (Eds.) *The Oxford handbook of moral psychology*, 47-71.

Darwin, C. (1887). Letter to William Graham. in F. Darwin (ed.). *The life and letters of Charles Darwin including an autobiographical note.* Vol. 1. London: John Murray. 315-317.

Dawkins, R. (2006) *The selfish gene* (3rd Ed.). New York, NY: Oxford University Press.

De Cruz, H., & De Smedt, J. (2012). Evolved cognitive biases and the epistemic status of scientific beliefs. *Philosophical studies*, 157(3), 411-429.

De Lazari-Radek, K., & Singer, P. (2012). The objectivity of ethics and the unity of practical reason. *Ethics*, 123(1), 9-31.

Dwyer, S. (1999). Moral competence. In K. Murasugi & R. Stainton (Eds.), *Philosophy and linguistics* (pp. 169-190). Boulder, CO: Westview Press.

Dwyer, S. (2006). How good is the linguistic analogy? In P. Carruthers, S. Laurence, & S. Stich (Eds.). *The innate mind: Culture and cognition* (pp. 237-257). Oxford: Oxford University Press.

Dwyer, S. (2007) How not to argue that morality isn't innate: Comment on Prinz. In W. Sinnott-Armstrong (Ed.), *Moral psychology* (Vol. 1). Cambridge, MA: MIT Press. 894-914.

Enoch, D. (2010). The epistemological challenge to meta-normative realism: how best to understand it, and how to cope with it. *Philosophical studies*, 148 (3), 413-438.

Enoch, D. (2011). *Taking morality seriously: A defense of robust realism.* Oxford University Press.

FitzPatrick, W. J. (2013). Debunking evolutionary debunking of ethical realism. *Philosophical studies*, 1-22.

FitzPatrick, W. J. (2014). Evolutionary theory and morality: Why the science doesn't settle the philosophical questions. *Philosophic exchange*, 44(1), 2.

Foot, P. (2001). *Natural goodness.* Oxford: Oxford University Press.

Fraser, B. J. (2014). Evolutionary debunking arguments and the reliability of moral cognition. *Philosophical studies*, 168(2), 457-473.

Gabennesch, H. (1990). The perception of social conventionality by children and adults. *Child development*, 61(6), 2047-2059.

Gert, B. (2005). *Morality: Its nature and justification*, New York: Oxford University Press.

Gill, M. B. (2009). Indeterminacy and variability in meta-ethics. *Philosophical studies*, 145(2), 215-234.

Gintis, H. (2007). The evolution of private property. *Journal of economic behavior & organization*, 64(1), 1-16.

Goodwin, G. P., & Darley, J. M. (2008). The psychology of meta-ethics: Exploring objectivism. *Cognition*, 106(3), 1339-1366.

Goodwin, G. P., & Darley, J. M. (2012). Why are some moral beliefs perceived to be more objective than others?. *Journal of experimental social psychology*, 48(1), 250-256.

Greene, J. D. (2007). The secret joke of Kant's soul. *Moral psychology: Historical and contemporary readings*, 359-372.

Griffiths, P., & Wilkins, J. (2010). When do evolutionary explanations of belief debunk belief?. In P. Sloan (ed.), *Darwin in the 21st century: Nature, humanity, and god.* Notre Dame, IN: Notre Dame University Press.

Haidt, J., Koller, S. H., & Dias, M. G. (1993). Affect, culture, and morality, or is it wrong to eat your dog?. *Journal of personality and social psychology*, 65(4), 613.

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion.* New York: Pantheon.

Hare, R. M. (1981). *Moral thinking: Its levels, method, and point.* New York: Oxford University Press.

Haselton, M. G., & Nettle, D. (2006). The paranoid optimist: An integrative evolutionary model of cognitive biases. *Personality and social psychology review*, 10(1), 47-66.

Hauser, M. D. (2006). *Moral minds: How nature designed our universal sense of right and wrong.* New York: Ecco.

Huebner, B., Lee, J. J., & Hauser, M. D. (2010). The moral-conventional distinction in mature moral competence. *Journal of cognition and culture*, 10(1), 1-26.

Hume, D. (2007). *A treatise of human nature: A critical edition.* In D.F. Norton and M.J. Norton (Eds.). Oxford: Clarendon Press.

Jong, J., & Visala, A. (2014). Evolutionary debunking arguments against theism, reconsidered. *International journal for philosophy of religion*, 1-16.

Joyce, R. (2006). *The evolution of morality.* Cambridge, MA: MIT PRess.

Joyce, R. (2013). Irrealism and the genealogy of morals. *Ratio*, 26(4), 351-372.

Joyce, R. (2014). Evolution, truth-tracking, and moral skepticism. In B. Reichardt (Ed.) *Problems of goodness: New essays on metaethics.* New York: Routledge.

Kelly, D., Stich, S., Haley, K. J., Eng, S. J., & Fessler, D. M. (2007). Harm, affect, and the moral/conventional distinction. *Mind and language*, 22(2), 117-131.

Kahane, G. (2011). Evolutionary debunking arguments. *Noûs*, 45(1), 103-125.

Kahane, G. (2014). Evolution and impartiality. *Ethics*, 124(2), 327.

Korsgaard, C. M. (1986). The right to lie: Kant on dealing with evil. *Philosophy and public affairs*, 325-349.

Ladyman, J. A. C., Presnell, S. M., Short, A. J. & Groisman, B. (2007). *Every thing must go: Metaphysics naturalized.* Oxford: Oxford University Press.

Machery, E. (2013). Morality did not evolve. Boston Colloquium for Philosophy of Science. Lecture conducted from Boston University, Boston, MA.

Machery, E. (2012). Delineating the moral domain. *Baltic international yearbook of cognition, logic and communication*, 7(1), 6.

Machery, E., & Mallon, R. (2010). Evolution of morality. In J. M. Doris (Eds.) *The moral psychology handbook* (pp. 3-47). New York, NY: Oxford University Press.

Mason, K. (2010). Debunking arguments and the genealogy of religion and morality. *Philosophy compass*, 5(9), 770-778.

Maynard Smith, J., & Parker, G. A. (1976). The logic of asymmetric contests. *Animal behaviour*, 24(1), 159-175.

Mikhail, J. (2000). Rawls' linguistic analogy: A study of the 'generative grammar' model of moral theory described by John Rawls in 'A Theory of Justice.' (Unpublished doctoral dissertation). Ithaca, NY: Cornell University.

Miller, G.F., (2007) Sexual selection for moral virtues. *The Quarterly review of biology*, 82(2)

Moll, J., Zahn, R., de Oliveira-Souza, R., Krueger, F., & Grafman, J. (2005). The neural basis of human moral cognition. *Nature reviews neuroscience*, 6(10), 799-809.

Neurath, O. (1983). *Philosophical Papers.* 1913-1946, Robert S. Cohen and Marie Neurath (Eds.) (Dordrecht: Reidel, 1983)

Nichols, S. (2004). After objectivity: An empirical study of moral judgment. *Philosophical psychology*, 17(1), 3-26.

Nichols, S. (2004). *Sentimental rules: On the natural foundations of moral judgment.* Oxford: Oxford University Press.

Nisan, M. (1987). Moral norms and social conventions: A cross-cultural comparison. *Developmental psychology*, 23(5), 719.

Nucci, L. P. (2001). *Education in the moral domain.* Cambridge, UK: Cambridge University Press.

Nucci, L. & Nucci, M. (1982). Children's social interactions in the context of moral and conventional transgressions. *Child development*, (53), 403-412.

Nucci, L. & Turiel, E. (1978). Social interactions and the development of social concepts in preschool children. *Child development*, (49), 400-407.

Nucci, L., Turiel, E., & Encarnacion-Gawrych, G. (1983). Children's social interactions and social concepts analyses of morality and convention in the Virgin Islands. *Journal of cross-cultural psychology*, 14(4), 469-487.

Peters, U. (2012). Evolution, moral justification, and moral realism. *Rivista italiana di filosofia analitica-junior*, 3(1), 8-18.

Plantinga, A. (2009). The evolutionary argument against naturalism: An initial statement of the argument. In M. Ruse (ed.) *Philosophy after Darwin: Classic and contemporary readings*. Princeton: Princeton University Press. 301-309.

Prinz, J. (2009). Against moral nativism. In D. Murphy and M. Bishop (Eds.). *Stich and his critics*. Oxford: Wiley-Blackwell. 167-189.

Putnam, H., 1975, *Mathematics, matter and method*. Cambridge: Cambridge University Press.

Quine, W. V. (1948). On what there is. *The Review of metaphysics*, 2(1), 21-38.

Rosas, A. (2012). Mistakes to avoid in attacking the moral/conventional distinction. *Baltic international yearbook of cognition, logic and communication*, 7(1), 8.

Russell, S. J. (1997). Rationality and intelligence. *Artificial intelligence*, 94(1), 57-77.

Sarkissian, H., Park, J., Tien, D., Wright, J. C., & Knobe, J. (2011). Folk moral relativism. *Mind and language*, 26(4), 482-505.

Sayre-McCord (2009). Moral realism. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Summer 2011 ed.). Retrieved from http://plato.stanford.edu/entries/moral-realism/

Schafer, K. (2010). Evolution and normative scepticism. *Australasian journal of philosophy*, 88(3), 471-488.

Shafer-Landau, R. (2012). Evolutionary debunking, moral realism and moral knowledge. *Journal of ethics and social psychology*, 7(1) 1-37.

Sinnott-Armstrong, W., & Wheatley, T. (2012). The disunity of morality and why it matters to philosophy. *The Monist*, 95(3), 355-377.

Skarsaune, K. O. (2011). Darwin and moral realism: survival of the iffiest. *Philosophical studies*, 152(2), 229-243.

Smetana, J., & Braeges, J.L. (1990). The development of toddlers moral and conventional judgements. *Merrill-Palmer quarterly*, 36, 329-346.

Street, S. (2006). A Darwinian dilemma for realist theories of value. *Philosophical studies*, 127(1), 109-166.

Street, S. (2008). Reply to Copp: Naturalism, normativity, and the varieties of realism worth worrying about. *Philosophical issues*, 18(1), 207-228.

Talbott, W. J. (2014). How could a "blind" evolutionary process have made human moral beliefs sensitive to strongly universal, objective moral standards?. *Biology and philosophy*, 30(5), 691-708.

Toner, C. (2011). Evolution, naturalism, and the worthwhile: A critique of Richard Joyce's evolutionary debunking of morality. *Metaphilosophy*, 42(4), 520-546.

Turiel, E. (1983). *The development of social knowledge: Morality and convention.* Cambridge, U.K.: Cambridge University Press.

Turiel , E . (1979). Distinct conceptual and developmental domains: social convention and morality in Howe, H. & Keasey, C. (Eds.), *Nebraska symposium on motivation, 1977: Social cognitive development.* Lincoln: University of Nebraska Press.

Vavova, K. (2014). Debunking evolutionary debunking. *Oxford studies in metaethics*, 9, 76-101.

Visala, A. (2011). *Naturalism, theism, and the cognitive study of religion: religion explained?.* Farnham, UK: Ashgate Publishing, Ltd.

Weigel, R. M. (1984). The application of evolutionary models to the study of decisions made by children during object possession conflicts. *Ethology and sociobiology*, 5(4), 229-238.

Wielenberg, E. J. (2010). On the evolutionary debunking of morality. *Ethics*,120(3), 441-464.

Wigner, E. P. (1960). The unreasonable effectiveness of mathematics in the natural sciences. Richard courant lecture in mathematical sciences delivered at New York University, May 11, 1959. *Communications on pure and applied mathematics*, 13(1), 1-14.

Wilkins, J. S., & Griffiths, P. E. (2012). Evolutionary debunking arguments in three domains. *A New Science of Religion*, 23, 133-146.

Wilson, J. Q. 1993. *The moral sense.* New York: The Free Press.

Wright, J. C., Grandjean, P. T., & McWhite, C. B. (2013). The meta-ethical grounding of our moral beliefs: Evidence for meta-ethical pluralism. *Philosophical psychology*, 26(3), 336-361.