Are moral judgements adaptations? Three reasons why it is so difficult to tell

Thomas Pölzler

To cite this article: Thomas Pölzler (2017) Are moral judgements adaptations? Three reasons why it is so difficult to tell, South African Journal of Philosophy, 36:3, 425-439, DOI: 10.1080/02580136.2017.1333354

To link to this article: http://dx.doi.org/10.1080/02580136.2017.1333354
Are moral judgements adaptations? Three reasons why it is so difficult to tell

Thomas Pölzler
Department of Philosophy, University of Graz, Austria
thomas.poelzler@uni-graz.at

An increasing number of scholars argue that moral judgements are adaptations, i.e. that they have been shaped by natural selection. Is this hypothesis true? In this paper I shall not attempt to answer this important question. Rather, I pursue the more modest aim of pointing out three difficulties that anybody who sets out to determine the adaptedness of moral judgments should be aware of (though some so far have not been aware of). First, the hypothesis that moral judgements are adaptations has been advocated in various different specificities and scopes, and on various different levels. Second, the three kinds of evidence that have most often been appealed to by discussants of this hypothesis require additional arguments. And third, there is significant reasonable disagreement about what moral judgements essentially are.

Introduction

Why are humans more prepared to fear snakes than guns? Why do sweet and fatty foods taste so good? Why do men all over the world prefer women with symmetrical faces, clear skin, full breasts and a low waist-to-hip ratio? An increasing number of scholars believe that just like many parts of the human body, a considerable number of our psychological traits are adaptations too. The reason for our having these traits is that in the environment inhabited by our ancestors, they positively affected reproductive success (e.g. Tooby and Cosmides 2005; Buss 2008).

In recent years such an evolutionary perspective has been increasingly applied to morality as well. Michael Ruse, for example, boldly proclaims: “morality…is…an adaptation. It is an adaptation like hands and teeth and penises and vaginas” (2009, 313). And according to Richard Joyce, “morality…can be given an adaptive explanation in genetic terms: …the present-day existence of the trait is to be explained by reference to a genotype having granted ancestors reproductive advantage” (2007, 2).

Adaptive explanations have been proposed for different aspects of morality (see Fraser 2010; James 2011). Some scholars, for example, have focused on moral behaviour (Wilson 1975); others on moral norms (Boyd and Richerson 2005), or moral sentiments or emotions (Frank 1988; Fiske 2010). In this paper I will be exclusively concerned with the hypothesis that moral judgements are adaptations, i.e. that thinking that (particular) things are right, wrong, good, bad etc. increased our ancestors’ reproductive success (with the term “moral judgement” referring to whichever mental states such thoughts are constituted by, whether these states are beliefs or desires).1 Prominent recent defenders of this variant of the adaptation hypothesis include Marc Hauser (2006), Philip Kitcher (2005; 2006; 2011), Sharon Street (2006), as well as the abovementioned Michael Ruse (1998; 2009) and Richard Joyce (2007).

1 Some scholars have used the term “moral judgement” to refer exclusively to moral beliefs. Let me re-emphasise that this paper does not involve such a cognitivist assumption. I will henceforth understand “moral judgement” in a way that is compatible with these judgements being constituted by desires (qua states with a world-to-mind direction of fit) as well. My understanding accordingly conforms to how the term “moral judgement” is predominantly used in contemporary meta-ethics (e.g. Svavarsdottir 2006, 163), as well as to how it has been used by most proponents and discussants of the hypothesis that moral judgements are adaptations (e.g. Kitcher 2005, 175; Joyce 2007, 51–57).
The truth of the adaptation hypothesis would have significant implications for our self-understanding as humans (e.g., Korsgaard 2010, 1–2). Some scholars have argued that it would also be relevant to normative ethics and meta-ethics. With regard to normative ethics, for example, the adaptedness of moral judgements has been claimed to support views such as that humans ought to act from altruistic motives (Richards 1986), or that consequentialist ethical theories are superior to deontological ones (Singer 2005; Greene 2008) (see also, e.g., Rottschaefer and Martinsen 1990; Rottschaefer 1991; Casebeer 2003).3 In meta-ethics the adaptation hypothesis has recently mainly been appealed to in attempts to “debunk” morality, in particular, in attempts to show that we do not have any moral knowledge (e.g., Joyce 2007) or that there probably are no objective moral facts (e.g., Ruse and Wilson 1986; Ruse 1998; 2009; Street 2006) (for attempts to “vindicate” morality in evolutionary ways see, e.g., Casebeer 2003; Kitcher 2011).

So is the adaptation hypothesis true? Do moral judgements really “lie in human nature”, in the sense that they are adaptations to some past environment? In what follows I shall not attempt to answer this important question. Rather, I pursue the more modest aim of pointing out three difficulties that anybody who sets out to determine the adaptedness of moral judgments should be aware of (though some so far have not been aware of). First, the hypothesis that moral judgements are adaptations has been advocated in various different specificities and scopes, and on various different levels. Second, the three kinds of evidence that have most often been appealed to by discussants of this hypothesis require additional arguments. And third, there is significant reasonable disagreement about what moral judgements essentially are.

**Difficulty 1: Variants of the adaptation hypothesis**

In order to be able to test the hypothesis that moral judgements are adaptations, one first needs to have a clear understanding of its content. I take it that at least the most basic features of adaptations are rather uncontroversial. As the notion is commonly understood – and as it is understood here – a trait qualifies as an adaptation if and only if it is the product of a particular evolutionary process, namely the process of natural selection. And a trait qualifies as a product of natural selection if and only if (1) it is heritable, (2) there was variation with regard to it, and (3) it increased its carriers’ biological fitness, i.e. led them to, on average, pass on a higher proportion of their genes to the next generation than individuals who had alternative traits (e.g. Darwin 1859; Buss et al. 1998; Buss 2008).

A lack of clarity about the adaptation hypothesis mainly concerns the notion of a moral judgement. Below I will argue that one serious difficulty in determining the adaptedness of moral judgements arises from the fact that there is significant reasonable disagreement about what moral judgements essentially are (see Difficulty 3).3 But even if we set aside this controversial meta-ethical issue, the adaptation hypothesis’ explanandum still needs to be specified in various ways. The reason for this is that natural selection may be and has been claimed to explain various different aspects of moral judgements, and in various different ways. Most importantly, the adaptation hypothesis is ambiguous with regard to the following three features: (1) its specificity, (2) its scope, and (3) its level.

**Specificity**

Moral judgements are a kind of practically normative judgements (e.g. Birnbacher 2007, 12–19). They are a kind of normative judgement in that they concern how things ought to be (as opposed to how things actually are); and they are a kind of practically normative judgement in particular in that they concern how subjects ought to act (as opposed to what subjects ought to believe).

---

2 The most famous attempt to derive normative-ethical conclusions from evolutionary theory is Social Darwinism, which roughly is the view that we ought to let natural selection operate freely, thereby promoting the “higher development” of our species (e.g. Spencer 1874; 1892). I do not mention Social Darwinism in the main text because due to its obvious empirical and philosophical inadequacies – and in particular to its problematic political ramifications – it tends to cast an unjustly poor light on any kind of evolutionary ethics.

3 I postpone discussion of this difficulty until the end because in order to fully understand its implications one already needs to have some knowledge about variants of the adaptation hypothesis (Difficulty 1), and in particular about evidence in evolutionary psychology (Difficulty 2).
Some philosophers and biologists have acclaimed that all practically normative or normative judgements are adaptations. Philip Kitcher (2005, 172), for example, argues that it was reproductively advantageous for our ancestors to have a “capacity for articulating rules and using those rules to shape our wishes, plans, and intentions” (i.e. those aspects of human psychology which underlie action). Machery and Mallon (2010, 20) even extend this claim to our “capacity to grasp and apply norms” as a whole (see also Street 2006, 156, fn. 2; 2009). As moral judgements are a kind of normative judgement and practically normative judgement in particular, the truth of these hypotheses would mean that there is some sense in which specifically moral judgements can be considered adaptations as well (see Machery and Mallon 2010, 4).

Contrasting with the above interpretations, however, moral judgements have also been claimed to be functionally specialized. On this narrower version of the adaptation hypothesis, making moral judgements increased the biological fitness of our ancestors in ways that are distinct from how other kinds of normative or practically normative judgements (such as epistemic, aesthetic or prudential judgements) might have done so. As a result, scholars such as Richard Joyce (2007) and Michael Ruse (1998) argue that humans evolved a psychological mechanism which is dedicated to a specifically moral form of normative cognition; a mechanism that is distinct from the mechanisms which allow us to make other kinds of normative judgements.

Scope
Interrelated with the issue of specificity, a second important ambiguity of the adaptation hypothesis concerns its scope, i.e. the question of which moral judgements the hypothesis applies to. Some proponents of the hypothesis have claimed that it increased our ancestors’ reproductive success to make particular moral judgements. A prime example for such a judgement concerns incest among close genetic relatives. Given that children of incestuous relationships have a high risk for disorders, disabilities and death, various scholars have argued that humans developed an adaptive tendency to judge incest morally wrong (e.g. Westermarck 1921; Lieberman, Tooby and Cosmides 2003).

Other proponents of the adaptation hypothesis have focused on broader kinds of moral judgements. According to Joshua Greene (2008, 43, 59–60; see also e.g. Singer 2005), for example, our ancestors’ rate of survival and reproduction depended on the infliction or non-infliction of “personal” harm, i.e. harm which is bodily, suffered by (one) particular identifiable person(s), and not arising from deflecting an existing threat (such as the harm caused by injuring a member of one’s group in order to take away his food). There was far less opportunity to harm others in impersonal ways. Accordingly, Greene argues, humans only developed an adaptive psychological tendency to intuitively condemn the infliction of personal (but not of impersonal) harm.

Finally, the hypothesis that moral judgements are adaptations has also been held in the sense that natural selection has had some influence on all of our moral judgements. On this view (e.g. Ruse 1998; Street 2006; Joyce 2007), all of these judgements in some way increased our ancestors’ biological fitness – judgements about genetic non-relatives as well as judgements about relatives, judgements about personal harm as well as judgements about impersonal harm.

Level
A last important differentiation, again interrelated to those considered above, concerns the level of adaptive explanations of moral judgements (see Joyce 2007; 2013a; 2013b; Ayala 2009; FitzPatrick 2014). Some proponents of the adaptation hypothesis have claimed that it is an adaptation that humans tend to make moral judgements with particular contents, such as the judgement about incest mentioned above. Others, in contrast, have rather focused on our moral capacity, i.e. on our ability to make moral judgements. They believe that while it increased our ancestors’ biological fitness to have concepts such as “right”, “wrong”, “good”, or “bad”, natural selection did not necessarily influence which particular things these concepts are applied to. These applications may well be (partly) determined by individuals’ culture.

Both content and capacity-orientated versions of the adaptation hypothesis have been held in numerous more specific forms and combinations. For example, it has been claimed an adaptation that humans have broad “evaluative tendencies” that influence what they judge right, wrong, good,
bad, etc. (for example, a tendency to judge things that promote one’s survival to be good, Street 2006, 119); that humans tend to make moral judgements about particular subject matters such as harm/care, fairness/reciprocity, in-group/loyalty, authority/respect and purity/sanctity (Haidt and Joseph 2004; Haidt and Graham 2007; Haidt and Björklund 2008); or that they accept a small number of very general moral principles, the application of which is determined within critical periods by our social environment (e.g. Mikhail 2013; Hauser 2006).

Table 1. Main ambiguities of the hypothesis that moral judgements are adaptations. All categories are orthogonal to each other, i.e. any element of one category can be combined with any element of the others. Moreover, elements of the category “Level” are non-exclusive, i.e. it is consistent to claim that both the content of the judgements at issue and our capacity to make these judgements are adaptations.

<table>
<thead>
<tr>
<th>Specificity</th>
<th>Scope</th>
<th>Level</th>
</tr>
</thead>
<tbody>
<tr>
<td>normative judgements</td>
<td>particular moral judgements</td>
<td>content of judgements</td>
</tr>
<tr>
<td>practically normative judgements</td>
<td>kinds of moral judgements</td>
<td>capacity to make judgements</td>
</tr>
<tr>
<td>moral judgements</td>
<td>all moral judgements</td>
<td></td>
</tr>
</tbody>
</table>

In the preceding sub-sections, we learned that the adaptation hypothesis can be and has been held in numerous different versions, varying with its specificity, scope, and level (see Table 1). This variety of ways of claiming moral judgements to be adaptations often has not been sufficiently accounted for. Many discussants of the adaptation hypothesis have failed to explicitly state which version/s of the hypothesis they are concerned with (e.g. Ruse and Wilson 1986; Ruse 1998; 2009; Korsgaard 2010). Some have even tentatively shifted between different versions, thereby arguing fallaciously; or they have misinterpreted the hypotheses held by others (see the below example of Prinz).

If the adaptation hypothesis were equally plausible on any of its numerous different versions these ambiguities, confusions and misinterpretations with regard to the above distinctions would be of minor consequence. But this is not what one finds. Different versions of the hypothesis significantly vary in terms of their truth conditions, i.e. in order to support or undermine these versions different arguments are required. Some of these arguments are much more promising than others (e.g. Joyce 2008; 2013a; 2013b). As an example, consider a recent debate between Richard Joyce and Jesse Prinz.

According to Joyce, humans have evolved a “tendency to make moral judgements” (2007, 134; emphasis in original). In defending this version of the adaptation hypothesis Joyce appeals to the claim that this tendency is universal, i.e. that humans at all times and places have judged things to be morally right, wrong, good, bad, etc. (see Difficulty 2). Prinz objects to this argument by pointing out that moral judgements are not universal at all:

Joyce’s...argument for moral nativism is that the content of morality is similar across cultures...Against this inference,...the variation within these domains is absolutely dizzying. Some societies engage in headhunting, cannibalism and slavery; some societies tolerate grotesque inequity; some societies are nearly egalitarian, while others have rigid class hierarchies; some societies have strict moral rules governing the body and others are extraordinarily permissive (Prinz 2008, 221–222).

This objection clearly fails. Even if Prinz were right that humans do not universally judge any particular action morally right, wrong, good, bad, etc. (a claim that Joyce actually does not make), they may still all have a general tendency to make moral judgements, i.e. a tendency to judge any

4 Different versions of the adaptation hypothesis also significantly vary in terms of the plausibility of their having certain normative ethical or meta-ethical implications. Recall, for example, how some proponents of the adaptation hypothesis claim that all kinds of normative judgements are adaptations. On this interpretation of the hypothesis, anti-realist or sceptical conclusions that might be derived with regard to moral judgements threaten to extend to theoretically normative judgements as well (judgements about what we ought to believe). However, many evolutionary arguments against the possibility of moral knowledge or the existence of objective moral facts presuppose the (objective) truth of such theoretically normative judgements (see Machery and Mallon 2010, 19–20).
thing morally right, wrong, good, bad, etc. At least initially this latter universality claim seems much more plausible. And in fact, Prinz’s objection bolsters the hypothesis that there is a universal tendency to make moral judgements itself, for it entails that although certain exotic cultures are otherwise very different from us, their members nevertheless make moral judgements (as pointed out by Joyce 2008, 259–260; see also Fraser 2010).

Given that different versions of the adaptation hypothesis vary in terms of their plausibility, the fact that discussants often have not explicitly distinguished between these versions, have shifted between them or misinterpreted each other constitutes a serious difficulty in assessing this hypothesis. It means that much of the literature on the adaptedness of moral judgements is prone to leading one astray. The remedy for this difficulty is straightforward. Discussants of the adaptation hypothesis must be clear about and explicitly state which version of this hypothesis they are concerned with, and they must mind the distinctions explained in this section in interpreting the claims and arguments of others.

**Difficulty 2: Evidence in evolutionary psychology**

For almost any of the numerous adaptive explanations of moral judgements distinguished in the previous section, competing non-adaptive explanations have been proposed as well. Some scholars have denied that moral judgements are products of evolution at all. This hypothesis entails that moral judgements either have not changed over successive generations or are not heritable (such as, for example, the human capacities for chess or handwriting; Machery and Mallon 2010, 23). Others have granted that moral judgements are products of evolution, but have nevertheless denied that they are adaptations. Such non-adaptive evolutionary explanations can take two forms. First, one may argue that moral judgements are by-products, i.e. that they inevitably come along with or are made possible by certain non-moral adaptations (such as by our highly advanced general intellectual capabilities; Darwin 1871; Ayala 2009; Fitzpatrick 2014; or by our capacity for having certain kinds of emotions; Prinz 2007; 2009). And second, one may argue that moral judgements are “random noise,” i.e. that they arose from random evolutionary processes such as fitness-neutral mutations (Williams 1966).

How can we determine whether the adaptation hypothesis is better than any of the above competing explanations? Investigations of the evolution of psychological traits are notoriously intricate. Neither are such traits directly captured in fossil records, nor can their existence or form readily be inferred from such records or other physical traces. In line with the methodology of evolutionary psychology more generally (see e.g. Andrews, Gangestad and Matthews 2003; Schmitt and Pilcher 2004; Simpson and Campbell 2005; Machery forthcoming), both proponents and critics of the adaptation hypothesis have mainly appealed to three kinds of evidence in support of their view. They have argued that moral judgements are/are not (1) reliable, economic, efficient and specific solutions to an adaptive problem, (2) universal, and (3) developing in the face of impoverished stimuli.

Below, I will suggest that one problem with regard to many arguments of the above kinds is that their soundness significantly depends on the controversial issue of the nature of moral judgements (see Difficulty 3). The difficulty that I attempt to raise awareness for in this section, in contrast, is methodological rather than conceptual. Claims about whether moral judgements are reliable, economic, efficient and specific solutions to an adaptive problem, are universal and develop in the face of impoverished stimuli cannot by themselves provide significant evidence with regard to the adaptedness of moral judgements in one way or the other. They require additional evidence which may be difficult to obtain.

---

5 As suggested by the remedy proposed below, this difficulty is not intrinsic (i.e. grounded in the nature of moral judgements or of adaptations), but rather of a practical or dialectical kind.

6 In some contexts it is useful to distinguish between traits that inevitably come along with adaptations and traits that are only made possible by them. One common way of doing so is by calling traits of the former kind “spandrels” (Gould and Lewontin 1979), and reserving the label “by-product” for traits of the latter kind (see Fraser 2010).
According to the most common argument in favour of the adaptation hypothesis, moral judgements are to be considered adaptations because they appear to be the product of design (for general defences of this kind of arguments see Williams 1966; Tooby and Cosmides 1992; 2005; Confer et al. 2010).

Recall the psychological traits mentioned in the introduction to this paper. These traits have been regarded as adaptations partly because they appear to be reliable, economic, efficient and specific solutions to adaptive problems. Our fear of snakes, for example, helped our ancestors to avoid being bitten by snakes. Men’s preference for women with symmetrical faces, clear skin, full breasts and a low waist-to-hip ratio led them to reproduce with particularly fertile and healthy partners. Proponents of the functional analysis argument in favour of the adaptation hypothesis claim that moral judgements must be regarded as an adaptation because they were a reliable, economic, efficient and specific solution to an adaptive problem as well.

Most commonly, moral judgements are supposed to have contributed to solving the adaptive problem of helping, i.e. of benefitting other individuals, or advancing their interests. Helping in this sense is widely taken to have increased our ancestors’ biological fitness (for example, because the recipients of their help often shared copies of their own genes (“kin selection”, Hamilton 1964a; 1964b); or because the help was later returned, either by the addressee himself/herself (“direct reciprocity”, Trivers 1971) or by others (“indirect reciprocity”, Alexander 1987)). Proponents of the helping version of the functional analysis argument believe that the function of making moral judgements – or of judging helping morally good and non-helping morally bad in particular – was to lead our ancestors to help others more reliably. Michael Ruse, for example, writes that “[m]orality has been put there by natural selection in order to get us to work together socially or to cooperate” (2009, 307; see also Joyce 2007, 13–73; James 2011, 59–62).

How convincing are functional analysis arguments in favour of the adaptation hypothesis? In my view there are at least two reasons not to overstate these arguments’ force.

First, we have less than full knowledge about our ancestors’ adaptive problems at the time at which moral judgements supposedly evolved (i.e. tens or hundreds of thousands of years ago). This means that we might not be able to tell for sure whether moral judgements were/were not a reliable, economic, efficient and specific solution to any of these problems (see e.g. Gould 1997; 2000; Buller 2005; Richardson 2007). For example, selective pressures from direct reciprocity – I help you because you will help me later – likely only favoured helping behaviour in case our ancestors repeatedly interacted with each other; for only under this condition was an individual’s help likely to be returned. It is unclear, however, whether our ancestors really interacted with each other in a sufficiently repeated manner. At the time of the supposed evolution of moral judgements, they may have already lived in far less small and close-knit societies than proponents of the above helping version of the functional analysis argument assume (Machery and Mallon 2010, 26).

Second, and even more importantly, functional analysis cannot by itself establish a trait’s having been shaped by natural selection in the first place. This is because appearances of design may not only be explained in adaptive terms, but just as well by non-adaptive evolutionary and non-evolutionary hypotheses (see Tooby and Cosmides 2005, 28). As a prominent example, consider the female orgasm (Machery and Mallon 2010, 34). There is a natural story about how this trait contributed to solving an adaptive problem. Females who were able to have orgasms, this story goes, simply will have been more motivated to have sex. Actually, however, the reason why females have orgasms is more likely that orgasms were selected for in males, and that male and female foetuses take similar developmental paths. This means that the female orgasm is a by-product rather than an adaptation (e.g. Lloyd 2005; Puts and Dawood 2006).

---

7 Adaptive problems are problems that a species faced repeatedly in the environment in which it evolved, and that significantly affected individuals’ reproductive success (Tooby and Cosmides 2005, 21–22).

8 Stephen Gould and Richard Lewontin (1979) famously coined the term “just-so-stories” for such empirically unsupported models of the adaptive function of a trait; referring to the famous children’s book by Rudyard Kipling, in which one finds fantastic explanations of why zebras have stripes, elephants acquired trunks, and so on.
The possibility of non-adaptive explanations of appearances of design suggests an important qualification of functional analysis arguments in favour of the adaptation hypothesis. In order for the fact that moral judgements functioned to solve an adaptive problem to possibly support the adaptation hypothesis, it would also have to be shown that this fact is best explained by the hypothesis – better than by the hypotheses that moral judgements are by-products, random noise, or not products of evolution at all.

**Universality**

Many discussants of the adaptation hypothesis have argued that (particular) moral judgements are to be/are not to be considered adaptations because these judgements are/are not universal, i.e. are/are not made by almost all humans at almost all times and in almost all societies. Recall, for example, the argument by Richard Joyce addressed in regard to Difficulty 1. In attempting to establish that humans have an adaptive tendency to make moral judgements Joyce appeals to the following view:

Morality (by which I here mean the tendency to make moral judgements) exists in all human societies we have ever heard of. Moral precepts are mentioned in the Egyptian Book of the Dead and in the Mesopotamian epic of Gilgamesh...we find the physical traces of morality far back in the archaeological record, at least into the early Upper Palaeolithic...and perhaps far beyond...like language, it is ubiquitous and ancient (Joyce 2007, 134–135, original emphasis; see also Hauser 2006, 53).

As with Prinz’s abovementioned response, discussions of universality arguments have so far mainly focused on the question of whether moral judgements are in fact universal (e.g. Stich 2008). Even Joyce himself (2013a) at one point acknowledges, however, that universality arguments suffer from a more fundamental problem as well. The truth of their premises would not lend strong support to their conclusion in the first place.

First, if a trait turns out universal, this does not by itself support the hypothesis that it is an adaptation. Some traits that are universal are not products of evolution at all. For example, the fact that the overwhelming majority of humans wear clothes (Prinz 2007, 372) and believe that the sun rises every morning (Machery and Mallon 2010, 14) is simply to be explained by these traits being easy to acquire and by humans living in similar environments and facing similar challenges. Other universal traits are by-products or random noise rather than adaptations. The reason why all humans have navels, for example, is not that navels increased our ancestors’ biological fitness, but only that they necessarily accompany umbilical cords (Buss 2008, 41).

Second, if a trait turns out as not being universal, this does not by itself support the hypothesis that it falls short of being an adaptation either. The fact that particular individuals do not show a trait’s observable characteristics may as well be explained by the hypothesis that the trait is an adaptation that is sensitive to certain environmental inputs – inputs which are just not present in the case of these individuals. A prominent example for such an adaptation is provided by calluses. Many humans lack the observable characteristics of calluses. But this non-universality is not due to the fact that calluses are not adaptations (they are, having evolved to protect deeper layers of the skin). Rather, calluses only develop in the face of friction, and humans’ skin often is not sufficiently exposed to friction (Schmitt and Pilcher 2004, 644).

In sum, then, the claim that (particular) moral judgements are/are not universal cannot by itself support any conclusion about the adaptation hypothesis either. In order for universality arguments to work, it would also have to be shown that the truth or falsity of this hypothesis is the best explanation of moral judgements being/not being universal.

**Poverty of stimulus**

Following similar arguments in linguistics, some discussants of the adaptation hypothesis have recently argued that children develop (particular) moral judgements even though their learning environment does not provide the stimuli that would be necessary for them to do so; and that this fact supports the hypothesis that these judgements are adaptations. In making his case for the adaptedness of the human moral capacity, Richard Joyce, for example, appeals to children’s
supposed ability to distinguish moral from conventional judgements:

Young children...at a remarkably early age...are able to discriminate among different kinds of deontic rules. Most notably, their capacity to distinguish moral from conventional transgressions emerges as early as the third year...These results from developmental psychology strongly suggest that the tendency to make moral judgements is innate [i.e. is an adaptation, as Joyce understands the notion of innateness, see Joyce 2007, 2] (Joyce 2007, 136–137).

Critics of the adaptation hypothesis have typically responded by suggesting that moral judgements actually do not develop in a way that exceeds the information available in children's environments (e.g. Prinz 2007; Sterelny 2010). But there is a more fundamental problem with this third kind of argument regarding the adaptedness of moral judgements as well. As was again acknowledged by Joyce (2013a) later, even granting the truth of their premises, poverty of stimulus arguments cannot by themselves strongly support or undermine any hypothesis about this matter.

First, the fact that a trait develops in the face of impoverished stimuli provides some evidence for its coming “from within us” – its being innate (in one particular sense of the term) – rather than its being learned.9 The fact therefore supports the hypothesis that the trait emerged from processes of evolution. It does not warrant concluding, however, that the trait is a product of natural selection in particular. This is because for a trait to develop in the face of impoverished stimuli is also consistent with it being a by-product of reliably developing adaptations and with its being reliably developing random noise. Noam Chomsky (1986), for example, the most famous proponent of poverty of stimulus arguments in linguistics, claimed that although humans’ ability to acquire languages is innate, it is better regarded as a by-product of the increased size and complexity of the human brain than as an adaptation.

Second, poverty of stimulus arguments cannot by themselves provide significant evidence against versions of the adaptation hypothesis either. That a trait falls short of being an adaptation is only one possible explanation of the fact that the trait does not develop in the absence of certain informational inputs. Another explanation is that the trait is an adaptation whose development is simply contingent on these inputs. For example, although humans may have an adaptive psychological disposition to develop belief in God (say, because this belief enhanced cooperation, Alcorta and Sosis 2005), for this disposition to become manifest certain information from individuals’ social or cultural environment may nevertheless be inevitable.

In sum, analogously to the arguments considered before, facts about moral judgements developing in the face of impoverished stimuli cannot strongly support or undermine the adaptation hypothesis by themselves. Proponents of poverty of stimulus arguments must also show that these facts are best explained by moral judgements being/not being adaptations.

Let us take stock. Discussants of the adaptation hypothesis have mainly appealed to three kinds of evidence in support of their views: evidence about moral judgements being a reliable, economic, efficient and specific solution to an adaptive problem; evidence about their being universal; and evidence about their developing in the face of impoverished stimuli. In the preceding sub-sections I argued that these kinds of evidence cannot by themselves lend strong support to hypotheses about the adaptedness of moral judgements – neither to the hypothesis that moral judgements are adaptations, nor to the hypotheses that they are by-products, random noise or not products of evolution at all. In order for the above arguments to work, the relevant evidence must also be shown to be best explained by any of these hypotheses.

What does this additional requirement mean for the assessment of the adaptedness of moral judgements? At the current stage of research we are not warranted to rule out that the requirement can be met. However, meeting it will most certainly be difficult (which likely is why only a few scholars have so far seriously attempted to do so).10 Suppose, for example, one would like to prove

---

9 For a trait to be innate in the above sense means for it to develop completely or largely independently of environmental input (Ariew 2006).
10 This is not to deny the existence of occasional brief discussions of the merits of adaptive vis-à-vis non-adaptive explanations of evidence
that the universality of our moral capacity is better explained by the adaptation hypothesis than by a certain by-product hypothesis. The two most natural ways of doing so are: (1) showing that while the human moral capacity is universal, the non-moral adaptation that this capacity is supposed to be a by-product of is not universal; and (2) showing that this non-moral adaptation could not give rise to our moral capacity in the first place (Joyce 2013a). Both of these strategies tend to be controversial, though. Relevant non-moral adaptations (such as our highly advanced general intellectual capabilities or our capacity for having certain kinds of emotions) are typically universal themselves. And whether a certain non-moral adaptation could give rise to a specifically moral capacity significantly depends on the contested issue of the meaning of moral judgements (see Difficulty 3).

In view of the above difficulty I suggest that discussants of the adaptation hypothesis broaden their horizon. For one thing, they should increasingly consider potential kinds of evidence other than functional analysis, universality and development in the face of impoverished stimuli (such as, for example, cross-species comparisons; see e.g. Ruse 1998, 227–230; Joyce 2001, 138). For another thing, where scholars do rely on these most common kinds of evidence, they should also increase their efforts to determine what best explains their findings about appearances of design, universality and development in the face of impoverished stimuli.

Difficulty 3: The nature of moral judgements

Before one can properly investigate whether moral judgements are adaptations, one does not only need to specify the particular aspect of these judgements that one is concerned with (see Difficulty 1), but also what moral judgements are in the first place.

Above I already made some hints in this direction. In particular, I suggested that moral judgements are practically normative, i.e. that they concern how their addressees ought to act. This condition of moral judgements is widely accepted. However, it is of course only necessary, and not also sufficient. Judgements such as that it is forbidden to drive above 50 kilometres/hour on a particular street or that one does not talk with one’s mouth full concern how their addressees ought to act as well. Yet, they are judgements of law and etiquette rather than of morality. So what distinguishes moral judgements from other kinds of practically normative judgements?

Questions about the nature of moral judgements range among the most debated and most contested in contemporary meta-ethics. For example, while some meta-ethicists maintain that moral judgements are constituted by beliefs (e.g. Smith 1994), others argue that they are constituted by emotions, attitudes of approval or disapproval, or other kinds of desires (e.g. Ayer 1936; Blackburn 2000). While some meta-ethicists believe that moral judgements apply universally (e.g. Hare 1952), others argue that things can be right or wrong for some individuals or cultures, but not for others (e.g. Harman 1996). There is disagreement about whether moral judgements entail categorical reasons for action (e.g. Kant 1785 vs. Foot 1972), about whether they are objective (e.g. Brink 1989 vs. Scanlon 1998), and about numerous other issues.

To be sure, this plurality of reasonable conceptual accounts of moral judgements does not complicate assessments of the adaptation hypothesis by itself. Important findings about the adaptedness of moral judgements could hold up on (almost) any of these different accounts, so that it ultimately does not matter which of them one assumes. In the real world, however, this possibility largely fails to obtain. Almost all arguments that have so far been put forward by discussants of the adaptation hypothesis are highly contingent on accounts of the meaning of moral judgements.11

Take, for example, functional analysis. Any plausible functional analysis argument entails that moral judgements motivated our ancestors to engage more reliably in certain fitness-increasing
behaviours (for example, in helping others). This raises the contested question of moral judgements’ conceptual relation to motivation (Björklund et al. 2012). According to motivational internalists, moral judgements are intrinsically motivating. In order for a person to possibly qualify as making such a judgement, s/he must have at least some (defeasible) motivation to act in conformity with his/her judgement (e.g. Smith 1994; Blackburn 2000). Motivational externalists, in contrast, deny any entailment of motivation (e.g. Brink 1989). Whatever one’s stance on this conceptual issue, it clearly affects how one assesses functional analysis arguments. In particular, arguments of this kind are ceteris paribus more plausible on motivational internalism than on externalism; for only on internalism does the fact that our ancestors judged a certain fitness-increasing behaviour morally right by itself guarantee that they were (defeasibly) motivated to engage in that behaviour.

The plausibility of universality arguments varies with conceptual accounts of moral judgements as well. Recall, for example, Richard Joyce’s claim that both the Egyptian Book of the Dead and the Mesopotamian Gilgamesh epic involve moral judgements. In Joyce’s view, moral judgements are highly specific mental states. For example, he requires that such judgements (qua speech acts) typically function to express both certain kinds of beliefs and certain kinds of desires, that they entail categorical reasons for action, that they counteract egoism, and so on (2007, 70–71). Several scholars have objected that assuming such a narrow account, we simply do not know whether ancient Egyptians and Mesopotamians made moral judgements (Stich 2008; Machery and Mallon 2010, 30–31; see also Joyce 2013a; 2013b). On broader accounts, in contrast – such as the view that moral judgements are constituted by states of approval or disapproval – it is significantly more plausible to attribute moral judgements to these ancient societies; and it is significantly more likely that moral judgements turn out universal in general (Machery and Mallon 2010, 31).

Finally, one’s understanding of moral judgements also influences the soundness of poverty of stimulus arguments. In the context of one such argument, we found Joyce claiming that even though young children’s environment lacks relevant information, they are able to distinguish between moral and conventional judgements. This claim is well-grounded as long as one defines moral judgements as judgements which are serious, universal, authority-independent, and harm- and fairness-related – because this is the particular account of moral judgements underlying most psychological research on the moral/conventional distinction (as particularly championed by Elliot Turiel (e.g. 1983)). In contrast, there is hardly any evidence about whether young children can distinguish moral from conventional judgements on alternative understandings of morality. On at least some alternative understandings, attributions of this ability actually seem less plausible. If moral judgements are defined as both relative and authority-dependent (see e.g. Kelly et al. 2007), for example, then they are more similar to conventional judgements than on Turiel’s account. It should therefore be more difficult for children to distinguish between moral and conventional judgements.

Considering the above examples, it is safe to conclude that research on the adaptation hypothesis does not only presuppose particular accounts of the nature of moral judgements, but that these accounts also significantly affect the findings that one makes. On some accounts it is much more likely that moral judgements turn out to be adaptations than on others. The difficulty arising from this conceptual contingency is obvious, and appears almost insurmountable. In order to possibly arrive at a definitive answer to the question of the adaptedness of moral judgements, one first seems to have to know what these judgements essentially are. But how can one realistically hope to achieve such knowledge, given that (as pointed out above) meta-ethicists have pondered over this question for centuries, and still have not come up with any generally agreed answer?13 In fact, however, the situation may not be just as grave as it seems (Pölzler forthcoming).

---

12 Joyce does not regard these features as necessary. Rather, he argues, that judgements must exhibit a sufficient number of them in order to qualify as moral (2007, 71).

13 The adaptation hypothesis’ contingency on controversial conceptual matters is particularly problematic for those who have attempted to derive meta-ethical conclusions from this hypothesis. Often these conclusions consist in or entail claims about the meaning of moral judgements. For example, the adaptedness of moral judgements has been thought to support that moral judgements are constituted by emotions (Kitcher 2005, 175); or that although they are constituted by beliefs that purport to represent objective moral facts, no such facts exist (e.g. Ruse 1998; 2009). But if it is true that one’s stance about the meaning of moral judgements affects one’s assessment of the adaptation hypothesis, then this hypothesis may fail to be theoretically neutral with regard to claims about the meaning of these judgements, and therefore cannot ground any argument in favour of these claims (Pölzler forthcoming).
One way of dealing with the adaptation hypothesis’ contingency on conceptual considerations is to go conditional. Scholars can show that if one defines moral judgements in a certain way, then these judgements are/are not likely to be adaptations (see Nadelhoffer and Nahmias 2007, 134). But even non-conditional progress may be more feasible than the above considerations suggest. Given that questions about the meaning of moral judgements are so controversial, we probably should give up on the idea of a strict logical priority of the conceptual over the empirical (see Brax 2009, 4, 11; Levy 2011, W1). Just as conceptual considerations can reasonably lead one to doubt certain empirical findings, such findings may provide reasons for revising one’s account of the meaning of moral judgements too. Suppose, for example, many judgements that are pre-theoretically widely and strongly held to be moral were found to be a reliable, economic, efficient and specific solution to an adaptive problem. Under certain circumstances this result may well support the conceptual claim that moral judgements are intrinsically motivating (see Gibbard 1990).

There are other prima facie plausible ways of alleviating the conceptual contingency problem as well (see e.g. Prinz 2015). In any case our considerations in this section have shown that in order to move forward on the question of the adaptedness of moral judgements, discussants need not only investigate empirical and methodological facts about these judgements showing appearances of design, being universal, developing in the face of impoverished stimuli, etc. It is just as important that one is aware of the conceptual assumptions implicit in such research, and critically examines these assumptions.

Conclusion
This paper obviously did not suggest that moral judgements are adaptations. However, it did not argue against this hypothesis either. All that I attempted to provide was an explanation of three difficulties which typically arise in investigating the adaptedness of moral judgements. First, the adaptation hypothesis has been advocated in various different specificities and scopes, and on various different levels. Second, the three kinds of evidence that have most often been appealed to by discussants of this hypothesis require additional arguments. And third, there is significant reasonable disagreement about what moral judgements essentially are.

Our investigation showed that many scholars have not so far been sufficiently aware of these difficulties. They have conflated different versions of the adaptation hypothesis, have drawn hasty inferences from the available evidence, have uncritically assumed certain conceptual claims about moral judgements, etc. At this point we therefore cannot tell for sure whether certain versions of the adaptation hypothesis are true (which means that the hypothesis accordingly provides a less solid basis for arguments in normative ethics and metaethics than has recently been assumed). Hopefully, however, papers such as this one can lead to greater conceptual and methodological clarity, and contribute to moving the debate about the adaptedness of moral judgements forward.

ORCID
Thomas Pölzler http://orcid.org/0000-0002-4311-0897

References


Received 8 March 2016, revised 9 May 2017, accepted 18 May 2017