

Chapter 6

Where Do Morals Come From? – A Plea for a Cultural Approach

Jesse J. Prinz

In recent years, there has been an empirical turn in ethics. Using the methods of psychology, neuroscience, behavioral economics, and evolutionary modeling, we have been able to make progress on old philosophical questions about the nature of morality. For example, much recent research has lent support to the view that emotions are integral to moral judgment. Unsurprisingly, empirical research in ethics has tended to be reductionist: the loftiest aspects of human behavior have been related to simple mechanisms that can be identified in the brain. The implicated mechanisms, most notably emotion circuits, are also known to have homologues in other creatures. This fact, together with evolutionary theory and behavioral ethology, has helped promote the idea that there is an innate moral sense. Nativist accounts have always been popular in cognitive science, so this outcome can hardly be surprising. But we should be cautious about importing that approach into the moral domain. Moral diversity within human populations suggests that, at the very least, culture is an important variable in shaping morality, and it is a variable that we cannot afford to overlook.

My goal here is to make a plea for a cultural approach to empirical ethics. I will begin by reviewing what I take to be the main empirical lessons about how we make moral judgments. Then I will argue that judgments, so understood, are not universal in content. This will lead to a discussion of where moral judgments originate. The brief answer is that cultural factors, unfolding across time, are crucial for understanding the content of morality. This has implications for how to think about the biological contributions to morality and the processes by which moral values are acquired.

J.J. Prinz (✉)
Philosophy Program at the Graduate Center,
City University of New York, New York, USA
e-mail: jesse@subcortex.com

6.1 What Is Morality?

6.1.1 *Emotion and Moral Judgment*

In order to understand from where morals arise, we need to know what morals are. By morals, I mean moral values. Values are long-standing evaluative attitudes or beliefs about what is good and bad. We evaluate many kinds of things: art, attire, wine, food, friends, manners, athletic performances, and so on. Typically, we evaluate things against standards, which include ideal features or exemplars, on the positive side, and objectionable features or exemplars on the other. To call something good or bad is usually to comment on its distance from a stored conceptualization of good-making or bad-making criteria or cases. For example, a wine might be judged as good if it has a balance of acidity and sweetness. In this respect, evaluative classification is like categorization more generally; it involves some kind of matching process. But there is also a crucial difference between evaluation and categorization.

To see this, notice that a person could taste a glass of wine and recognize it as such, without having any view about whether it is good. One can even discern a balance of acidity and sweetness without judging that this balance is good. Judging that such balance is good requires a *response* to it. To qualify as a positive evaluation, the response has to have a motivational force; it has to promote consumption of the wine. When we evaluate things positively, we are usually thereby attracted to them. Negative evaluation, in contrast, motivates avoidance, cessation, or withdrawal.

If evaluations are responses to recognized features, and those responses have motivational force, then it is natural to suppose that evaluations are *emotional* in nature. Emotions are responses to things that go beyond recognition, and emotions promote various forms of approach and avoidance. To evaluate a wine as good, it is plausible that the wine causes a positive emotion in us: a kind of pleasure. Alternatively, we might say that a wine is good without experiencing such pleasure. For example, we might suppose that a wine is good because the sommelier recommended it. But in such cases, our evaluations are deferential, or parasitic on another evaluator. The sommelier, we can presume, takes pleasure in good wine, or has at least mastered a list of preferences from someone whose pleasure is regarded as authoritative.

I think such emotional responses are the mark of the evaluative. Without emotional reactions, we can categorize, but we cannot appraise things as good or bad. A dispassionate appraisal is possible only by deference to a passionate judge. In philosophical jargon, such an appraisal would be a case of “mentioning” rather than “use.”

Against this background, it is plausible to suppose that moral evaluations are also emotional. To judge that infanticide is bad is not just to say that it involves a certain activity (the intentional killing of a baby), but also to find the activity abhorrent. This simple observation lies behind a philosophical tradition called sentimentalism according to which moral values are sentiments (prominent defenders include Hume

1739; Smith 1759; Ayer 1952; Blackburn 1984). A sentiment can be defined as a disposition to have an emotional response. Thus, to have the value that infanticide is bad is to have the disposition to have an emotional response (of a kind to be described below) towards killing babies. Moral judgments are occurrent emotional states towards actions, and moral values are dispositions to make such emotion-laden judgments.

The sentimentalist tradition in philosophy has gained renewed support from cognitive science. Over the last 15 years, there have been numerous empirical studies investigating what goes on when people make moral judgments. These studies have varied tremendously in design and methodology, but they have converged on the conclusion that emotions are centrally involved in moral judgment (for a recent overview see: *Emotion Review*, 2011, volume 3). Neuroimaging studies have shown that emotion centers of the brain are active when people consider moral dilemmas (Greene et al. 2001), read sentences describing moral violations (Moll 2002a), view morally significant pictures (Moll et al. 2002b), or encounter morally questionable playing partners in economic games (Sanfey et al. 2003). Behavioral studies have shown that emotion induction causally influences moral judgments. For example, people make more severe judgments of wrongness when situated at a dirty desk or when smelling noxious odors (Schnall et al. 2008), and when they experience hypnotically induced disgust. Induction of anger through films or autobiographical recall can also lead to harsher judgments (Lerner et al. 1998), and induction of happiness can lead people to be more utilitarian in orientation, approving the violent sacrifice of one innocent person to save five people in danger (Valdesolo and DeSteno 2006). Working with collaborators, I have sought to replicate and extend these findings. We have shown that disgusting beverages make moral judgments harsher (Eskine et al. 2011), and that irritating music increases negative moral judgments, and uplifting music increases positive moral judgments (Seidel and Prinz 2013). All this suggests that people use emotions as information when they decide whether something is right or wrong: when asked to make a moral evaluation, people introspect and report the intensity of their feelings.

It also has been shown that emotions can lead people to make moral evaluations even when they can't produce reasons to justify those evaluations (Haidt 2001). In a pilot study on this theme, I was able to show that people harshly judge a child molester even when his victim is unharmed and has no way of recalling or being traumatized by the incident (Prinz in press). Such findings suggest that we report our moral values by introspecting on our emotional states. The degree of negative emotionality determines our assessment that something is morally bad, even in the absence of supporting reasons. This suggests that emotions are *sufficient* for evaluating something as bad.

Emotions may also be *necessary*. Individuals who have impairments in emotional responsiveness show corresponding impairments in morality. Criminal psychopaths, for example, show deficits in negative emotions, and also seem to treat moral rules as mere social conventions (Campagna and Harter 1975; Blair 1995). Individuals with frontotemporal dementia suffer from a diminished capacity to evoke emotional states and show a corresponding tendency to see morals as

conventional (Mendez et al. 2005). Such findings suggest that, absent certain emotions, we lose the capacity to make moral judgments. The personal *evaluation* that something is morally bad gets replaced by the social *categorization* that something is prohibited by the community.

These empirical results can be systematized by the sentimentalist theory of morality. Emotions seem to be sufficient and necessary for moral judgments, and that can be explained by assuming they are component parts of such judgments. The judgment that something is wrong *consists in* a negative feeling toward it. If negative feelings are introduced extraneously (e.g., by noxious smells), we will feel more intense emotions and report that we think things are more wrong than we would report under other conditions. If emotional responsiveness is diminished, things seem less wrong than they otherwise would.

This story about moral judgments can be extended to other kinds of evaluations. For example, recent neuroimaging studies suggest that emotions are involved in aesthetic judgments (Kawabata and Zeki 2004; Vartanian and Goel 2004) and that reduced emotionality promotes aesthetic indifference (Chapman et al. 1976). This raises a question: what distinguishes moral judgments from other kinds of evaluative judgments?

The answer I favor is that moral judgments involve a distinctive class of emotions. It has been shown that other-directed moral judgments characteristically involve anger, contempt, or disgust and these are tuned to different kinds of transgressions (Rozin et al. 1999). We become angry about crimes against persons, contemptuous of crimes against the community, and disgusted by crimes against nature. There are also self-directed moral emotions, which may also have different functional roles. Guilt seems to arise when we harm another person, and shame arises when we do something that others might regard as unnatural or grotesque (Prinz, unpublished data). I have proposed that moral values are constituted by sentiments that dispose us to feel anger, contempt, or disgust towards others and guilt or shame towards oneself. To have a moral value requires the disposition to feel both these other-directed emotions and self-directed emotions. Sentiments involving different emotions, or lacking in both the other- and self-directed dispositions do not qualify as moral judgments. Aesthetic values, for example, involve different emotions, and drinking bad wine may cause disgust, but it won't cause guilt or shame.

The picture so far can be summarized by saying that moral values are sentiments, and sentiments are dispositions to feel both the self- and other-directed emotions of a certain kind. The emotions I have been discussing can be classified as emotions of *blame*, since they are socially directed and punitive in nature. A person who is the target of anger, contempt, or disgust will feel punished in virtue of being regarded in these negative ways, and each emotion will also motivate behaviors (such as aggression, in the case of anger, or avoidance in the case of disgust) that are tantamount to forms of punishment. Legally proscribed forms of punishment, such as torture, execution, banishment, and incarceration, can be seen as social inventions that institutionalize the kinds of actions we might be inclined to carry out given our emotions of blame. In equating moral values with sentiments, I mean to suggest that

they are sentiments and nothing more. Thus, beyond the cognitive representation needed to represent a certain type of action (e.g., stealing), sentiments are sufficient for regarding that action type as wrong; to think stealing is wrong consists in our negative sentiment towards it. One might come to have many cognitively represented beliefs about wrongdoing (e.g., that stealing decreases social stability or impedes with autonomy), but these are best described as contingent theories about what makes things wrong, which, unlike sentiments, are neither necessary nor sufficient for having moral values.

In addition to the punitive attitudes that I have been discussing, there are positive moral values that revolve around praise, rather than blame. Praise and blame play asymmetric roles in morality. For example, we rarely praise people for conforming to moral rules, but we do blame people for deviating. Praise is usually reserved for supererogatory acts, such as charity, or other forms of self-sacrifice. Positive emotions, such as gratitude and esteem, are likely to underwrite the values that lead us to appraise such acts as good, but I will not survey those emotions here. My focus will be on moral prohibitions since, given the asymmetry, these are the mainstay of moral life.

6.1.2 The Content of Morality

I have characterized moral norms in terms of the emotions that arise when we make moral judgments. It is by means of these emotions that we can identify when someone is moralizing, even if their values differ from our own. To that extent, the characterization is content neutral. It does not define morality by its subject matter. This is important because, as we will see, people moralize different things. Indeed, almost anything could be moralized. We moralize interpersonal actions, thoughts, character traits, personal habits, self-presentation, and so on. Even things outside our control can be regarded as morally wrong; consider the Christian doctrine of original sin. That said, the sentimentalist framework presented here can be used to make some broad generalizations about the content of morality. Such generalizations have already been hinted at with the taxonomy of other-directed emotions.

Recall that anger, contempt, and disgust arise in response to different kinds of transgressions. In particular, they vary as a function of who is victimized by a transgression: anger is a response to crimes against persons; contempt arises in response to crimes against community; and disgust responds to crimes against nature. These broad categories can be further refined by reflecting on ways that persons, community, and nature can be assailed against. Consider, first, crimes against persons. This category includes physical harm, as when a person is hurt, mutilated, or killed. But the category also includes violations of individual rights. Rights, in the Western tradition, are usually regarded as entitlements: the right to own property, to free speech, to education, to choose a religion, and so on. Preventing someone from having something to which she or he is entitled is usually regarded as a moral wrong; entitlement itself is usually understood as a moral, not just legal, construct.

When this happens, anger is the dominant emotional response. Anger can also arise in response to violations of distributive justice. If a distribution is unfair, those who get less than their share have been victimized. Thus, unfairness is a crime against persons, and it incites anger.

Contempt arises when a transgression is construed as an assault against the community. This happens, for example, when someone disrespects authority. Disrespect to authority can threaten to undermine the structure of the community, even if no one is directly harmed. The community can also be threatened when someone fails to conform to a social status hierarchy. Stepping out of line (e.g., looking down on one's parents or the elderly) is viewed with contempt. In addition, each social class tends to view the others with a degree of contempt, and this may serve to keep classes in their place. Contempt is also the emotion that arises when there is a transgression against public goods, such as vandalism or cases where a politician embezzles public funds or violates public trust. Here, again, the community as a whole is harmed.

Disgust is the response to unnatural acts. In non-secular societies, such acts are usually construed as crimes against God or gods (Shweder et al. 1997). Within a religious framework, supernatural agents are the authors and regulators of nature, so crimes against nature are forms of sacrilege. Secular societies continue to regard certain acts as unnatural, even if there is no obvious human victim. This is especially true of acts that involve the body. For example, some sexual behavior is considered immoral in many societies, such as bestiality, incest (even if consensual), and exhibitionism. There are also norms governing appropriate appearance (e.g., gender specific attire, broad conformity to current clothing styles, appropriately groomed hair, and cleanliness). Minor violations may provoke ridicule, but more extreme cases are likely to provoke disgust. In addition, there are norms governing diet. Kosher laws are a non-secular example, but secular dietary norms are also easy to find: some cultures prohibit consumption of horses, animals that have been domesticated as pets, and insects, for example. The consumption of human flesh, even if the person died naturally, is also widely condemned, and, in all these cases, the emotion of condemnation is disgust.

These examples illustrate two things. First, the content of morality is highly varied. Many moral values have little to do with harm, and every aspect of human life can be subject to moral rules. Second, in some broad, metaphorical sense, negative moral values can be regarded as concerning actions that are directed against one of three categories: persons, community, or nature. These categories may turn out to exhaust the moral domain (e.g., can there be crimes against abstract objects?). Each category is governed by a different moral emotion. We also have moral values pertaining to things other than actions, such as sinful thoughts or vicious character traits, but these attitudes may depend on a connection to actions: thoughts and traits potentially affect behavior. Thus negative moral values can be captured by the schema:

An agent A's doing/having/being X is bad iff by X, A (potentially or actually) has an effect on victim V, where V is construed as a person, a community, or nature, and, depending on that construal, an evaluator E who so construes A's X-ing will

have the corresponding emotion of blame (anger, contempt, or disgust, if E is a third party, and guilt or shame if E=A)

This schema summarizes the foregoing discussion. It gives us an account of what moral values are, and we can now reflect on where they come from.

6.2 Where Do Moral Values Come From?

6.2.1 Is Morality Innate?

Psychologists and cognitive neuroscientists typically assume that they are studying universal facts about human nature. Studies of memory, attention, and reasoning are presented as revealing the laws of thought, akin to natural laws in other sciences. A typical study of memory span, for example, rarely begins with the qualification that this is how memory works among American college students, or whoever makes up the subject pool. The demography of the subjects is (roughly) indicated, but it is presumed that demography has little impact on the results. The presumption rests on the view that these basic faculties of the mind are innate, and relatively unaffected by learning. There is, in other words, an implicit nativist bias in the way the sciences of the mind are typically pursued.

The nativist bias is also implicit in some of the empirical work on morality. Psychological and neuroimaging studies of moral cognition rarely look at culture as a variable (consider the citations in Sect. 6.1.1). This implicitly assumes that moralizing is part of the universal human bioprogram. Many of these studies say little about the *content* of our moral values and focus more on the processes involved in moralization. To that extent, they are neutral about the origins of our specific values, even if they are implicitly nativist about the mechanisms that allow moralization. Some other research, however, takes a stance on questions of content.

We can see that there are three basic positions one can take with respect to the innateness of morality:

- *Strong Nativism*: The content of our moral values is innately determined or strongly constrained.
- *Weak Nativism*: We have an innate faculty for acquiring moral values, but the content of those values is not strongly constrained.
- *Anti-Nativism*: We have no innate faculty dedicated to morality.

As I read the literature, Weak Nativism is often implicitly presumed, and Strong Nativism is sometimes explicitly defended. Anti-Nativism is a minority position, which is rarely implicitly or explicitly endorsed. I myself am a methodological anti-nativist, which means I think we should assume that a faculty is not innate until evidence leads us to say otherwise. In the case of morality, some researchers think the evidence supports Strong or Weak Nativism. I am not convinced. Here I will briefly consider some of the evidence (see also Prinz 2007a).

Let me begin with Strong Nativism. One research program that has a Strong Nativist orientation is the so-called moral grammar approach, which pursues an analogy between morality and language (Mikhail 2000; Hauser 2006a). It is ironic that defenders of this approach tend toward Strong Nativist positions; given that language is generally regarded as weakly innate (languages vary hugely in phonology and vocabulary). Officially, defenders of moral grammar say morality can vary too, but much of their research is designed to establish universal moral content. Notably, Mikhail and Hauser have acquired evidence that most people respond in predictable ways to a range of “trolley dilemmas,” in which an agent performs an action that leads to one person’s death in order to save five others. For example, most people think it is wrong to push someone into a runaway trolley’s path in order to save five people further down on the track, but it is permissible to divert a runaway trolley onto a track where it will hit one person instead of five. Responses to such dilemmas are cross-culturally robust, but people have great difficulty articulating the principles on which they are relying. Nativists interpret this as evidence for unconscious rules, analogous to those used in language processing.

Trolley experiments, however, can also be interpreted in other ways. The fact that people in different cultures give similar responses might be explained by prototype effects. When people learn the concept *murder*, the paradigm cases involve direct intentional physical assault, not indirect harms. The reason for may have nothing to do with innateness. All cultures must have rules to stop people from directly and intentionally aggressing against each other, on pain of societal collapse. Rules against indirect harms, however, are less prevalent, because there are fewer circumstances within a society when indirect actions will result in someone’s death, and a society that failed to have such rules might be relatively stable. The pushing scenario conforms most closely to the kind of actions that every society is likely to condemn. It is more clearly an instance of murder than the scenario in which a person is killed as the side-effect of diverting the trolley. In the “diversion” scenario, the death is also less salient and the cause of death for the one person is rendered comparable to the cause of death for the five, making the comparison between the two outcomes vivid. So there need not be any unconscious rules at work here. People are taught that murder is wrong by means of prototypical cases, and they tolerate killing more readily when it departs from the prototype, lacks salience, or is rendered comparable to an alternative action that involves the same kind of killing but greater losses.

Another research program that is committed to some degree of strong nativism is the moral domains theory of Turiel (1983). Turiel argues that genuine moral rules involve harms, and that other kinds of rules are mere conventions. In comparison to conventional rules, rules pertaining to harms are treated as more serious and less dependent on authority. Turiel believes that this pattern of conceptualization is innate. But there are five reasons for rejecting this position. First, harm norms are judged to be authority dependent in some studies (Kelly et al. 2007). Second, norms pertaining to diet, sexuality, and hierarchy are treated as equally serious by some groups (e.g., Vasquez et al. 2001, on Filipinos; Nisan 1987, on Palestinians). Third, there is a simple learning story available to explain why moral norms are treated

differently than conventional norms. Moral norms are taught by emotional conditioning, and once emotional attitudes have been internalized, the norms feel serious (i.e., emotionally evocative) and somewhat independent of authority (i.e., we are conditioned to feel emotions towards these acts even if we are in a community where others don't have such emotional dispositions). Fourth, there is massive cultural variation in attitudes towards harm. Many societies have practiced slavery, corporal punishment, judicial torture, agonizing body modification, blood sports, animal cruelty, spouse beatings, and virtually unconstrained brutality against out-groups; hardly evidence for an innate prohibition against harm. Finally, the fact that many societies do have moral norms against some forms of harm (notably gratuitous harm against the in-group) can be explained by the fact that we devise such prohibitions as a condition on societal cohesion. It does not take innate mechanisms to realize that tolerated killing will lead to social unrest. The fact that such norms have a highly moral status worldwide may also reflect the fact that anger is a natural response to aggression in the first-person case. Given that we are all disposed to get mad when others try to harm us, it is not surprising that the more general stricture against harm, which extends to third parties, is grounded in anger. This grounding helps give harm norms their moral cast.

Another research program that has a Strong Nativist flavor is evolutionary ethics. Evolutionary ethicists admit that nativism is compatible with moral diversity (e.g., Krebs 2008), but they tend to offer evolutionary models that emphasize highly predictable behaviors, suggesting that morality may be strongly constrained. Most of this work focuses on altruistic behaviors, in which individuals incur costs to benefit others. Models that use iterated economic games have shown that cooperative strategies, such as reciprocal exchanges, increase fitness, suggesting that cooperation may be an evolved response. The evolutionary interpretation gains support from the fact that general purpose reasoning, together with hyperbolic discounting, does not predict cooperation. Reasoning would lead people to see the value of defection, yet we do, in fact cooperate. Other prosocial behaviors, such as helping people in need and sharing resources, are also widely documented. Like cooperation, these behaviors are hard to explain by appeal to reasoning, which suggests that they may be innate. The evolutionary approach is bolstered by ethological research on non-human primates. Monkeys and apes are known to reciprocate, share, and help (de Waal 1996; Brosnan and de Waal 2003; Hauser et al. 2003). It is presumed that these behaviors are unlearned in our primate relatives and may reflect hard-wired precursors to our own prosocial tendencies.

There are several reasons to resist the evolutionary approach to morality. First, most of the work concerns moral behaviors, not moral judgment. By that, I mean behaviors that we now happen to regard as morally praiseworthy (cf. Joyce 2006, on this distinction). In principle, a species could evolve to act in ways we find praiseworthy without evolving a capacity to praise. That is, there can be moral conduct without moral judgments. This point is especially problematic when it comes to extrapolating from animal research, since most of that work concerns "altruistic behaviors," and not moral judgments per se. Moral judgments have two features that are unlikely to be found in many other species. First, they require a disposition for

self-directed emotions. Evidence for guilt and shame in non-human primates is scant at best. If apes get angry when conspecifics trespass against them, it does not follow that they would feel guilty for trespassing themselves. Reactive aggression is not the same as forming a moral judgment; self-directed dispositions are needed as well. Second, there is only a little anecdotal evidence that non-human primates have concern for third parties. Moral rules quantify over agents and action types. They are not restricted to the second-person. If apes get angry when conspecifics trespass against them, it does not follow that they would get angry if one conspecific trespassed against another, especially a non-relative. If they do not do this, then their anger reactions don't stem from values that have the schema indicated above.

A second problem with animal models is that there are profound differences between apes and humans. Chimps often fail to share with long-time companions, even when there is no cost (Vonk et al. 2008), and they are often highly aggressive in the wild. Goodall (1986) documents cases of chimpanzee warfare, calculated murder, infanticide, and cannibalism. Wrangham et al. (2006) report that chimpanzees are alarmingly violent; comparing several wild populations to a small-scale human group known for aggression, the found male chimps were 384 times more likely to engage in a violent attack than were their human counterparts. One might reply that apes simply having a different morality than ours, but given these differences, the burden is on the nativist to say why ape behavior must be interpreted as based on moral judgments, as opposed to some other kind of motivations. After all, not every kind human act is a result of morality (threat of punishment, instrumental gain, friendship are among other motivators). This is not to deny that some forms of ape altruism might have biological roots in common with our own, but only to emphasize that we must be cautious about over-attributing human-style moral tendencies to apes. There may be important discontinuities.

Moving beyond comparative research, evolutionary theorizing suffers from another limitation with respect to Strong Nativism. Evolutionary models have shown that it is difficult for altruistic behaviors towards non-kin to evolve through individual selection. If I mutate to reciprocate, but you do not, I will suffer a profound decrease in fitness. This has led to a widespread endorsement of group selection models. But group selection raises the possibility that widespread reciprocity evolves culturally, rather than biologically. Of course, nativists can offer alternative explanations that avoid group selection, but once such models are shown to be viable the pressure to explain altruism biologically decreases. More generally, there is something suspicious about any argument that moves from a demonstration of fitness enhancement to a conclusion about innateness. Many behaviors that would enhance fitness are not evolved; over generations, groups can learn to perform actions that are beneficial and avoid actions that are harmful. To show that morality is innate, models are not enough. Evidence must also show that specific moral rules are *universal* and *unlearnable*.

With respect to universality, evolutionary approaches tend to suffer from a dearth of empirical support. The models might be taken to suggest that all people are equally altruistic, but, in reality, there is considerable cultural variation. Sharing, for example, varies with respect to competing principles of distribution. In America, the

preferred principle is equity (distribution as a function of achievement), in China there is a preference for equality, and in India there is a preference for distribution as a function of need (Leung and Bond 1984; Berman et al. 1985). It is hard to think of sharing beyond one's kin as a biological norm given the rise of global capitalism, widespread opposition to taxation, and staggering discrepancies in wealth. Similar conclusions can be drawn about helping. Trivial, low-cost, helping behaviors, like picking up a pen that some has dropped, differ dramatically from place to place, with Rio residence coming out on top and New Yorkers bringing up the rear (Levine et al. 2001). Cultures also vary in the degree to which helping the needy is seen as a cultural requirement. In the United States, helping strangers with moderate neediness is considered entirely optional, but it is morally mandated in India (Miller et al. 1990). Americans, unlike Indians, also seem to think the obligation to help someone in moderate need depends on whether we like that person (Miller and Bersoff 1998). In general, we do amazingly little to help the needy. Preventable diseases claim about nine million lives a year, as does starvation, suggesting an annual toll that dwarfs the holocaust, and nearly universal crimes of omission.

Finally consider learnability. Evolutionary ethics presumes that we would not engage in prosocial behavior if we relied on domain-general resources such as reasoning. Given the human tendency to discount the future, we would behave unethically to reap short-term rewards. The fact that we are generally pretty good to each other is taken as evidence that morality is innate. Here again, one wants to distinguish moral behavior and moral attitudes. After all, squirrels are pretty good to each other, but no one thinks they have innate morality. But putting this issue aside, one can also deny the premise that domain general resources would not lead to cooperative behavior. It is true that reasoning might not be up to the task, but emotions are well suited to this purpose. Suppose I fail to cooperate with you and you get mad. I may be frightened of punishment or sad about losing you as a partner. Thus, your anger can condition me to associate negative emotions with defection. Suppose now there is an opportunity for me to defect without you finding out. Reason might lead me to do so, but emotions operate somewhat independently of reason, and my negative associations may promote cooperation even in this situation where free-riding is an option. Notice that this appeal to emotions as mechanisms of cooperation is also central to evolutionary models (Trivers 1971; Frank 1988). The point here is that once we recognize that emotions are the glue that promotes prosociality, there is actually less pressure to assume that morality is innate, because emotional dispositions can be easily learned through conditioning. Emotions may be evolved for selfish purposes (anger protects us against threats, and sadness makes us withdraw in times of loss), but selfless dispositions can arise when these selfish patterns are conditioned by interactions with others. Your rage becomes my loss, so I learn to avoid making you angry.

Expanding this last point, the acquisition of prosocial behavior needs two ingredients. First, if I defect in my dealings with you, you will get mad. That's not a moral response; it's just reactive aggression. Second, if you get mad, I feel bad and associate this with defection, leading to increased tendency to cooperate. These two steps could even be realized in non-human primates. Human beings may go on to a

third step: we generalize moral rules and apply them in cases where we have no direct involvement. This might be explained by the fact that human beings have two capacities that are underdeveloped in primates: imitation and abstract thought. Imitation leads us to mimic the reactive aggression of those who get mad at us. Abstraction leads us to internalize emotional dispositions in a way that can generalize across individuals, because we can represent actions abstractly rather than merely first-personally, as something I do. Thus, if you get mad at me for defecting, I might come to have bad feelings about defecting in general, whoever does it, and I might adopt your anger response when I encounter the defection of another. I don't want to suggest that this is the whole story. There may be innate behavioral tendencies that contribute to the moral rules with which we end up. But these simple observations suggest that the acquisition of moral rules need not involve any highly specialized mechanisms.

This last point allows us to move from Strong Nativism to Weak Nativism. Strong Nativists claim that the content of morality is innately determined or strongly constrained. I have tried to cast doubt on that conclusion by briefly reviewing some of the leading research programs that emphasize innate content. The content of moral rules is variable, and convergence can be explained without innateness. Now, with this simple story about psychological prerequisites to morality, we can see that even Weak Nativism may be mistaken. The acquisition of moral rules may not depend on any kind of morality acquisition device (Sripada and Stich 2005), but may instead derive from cognitive resources that evolved for other purposes (emotions, imitation, abstraction). Far more would need to be said to firmly establish that domain general resources are up to the task. For present purposes, I am content with the conclusion that we should be open to this possibility. Just as religion may arise in all cultures without a religion module, morality may be a byproduct of capacities that are not specific to the moral domain. As a methodological anti-nativist, I'd like to see more evidence for domain specificity before concluding that morality is even weakly innate.

6.2.2 Morality, Culture, and History

I just reviewed evidence for moral nativism and found it wanting. I also indicated some of the proximate psychological mechanisms that may be involved in the acquisition of moral rules. But what about more distal factors? Why do we have the rules that we do? If I am right, the answer to this question cannot be given solely by evolutionary theory, but must recruit the resources of cultural anthropology and history. The factors that give rise to moral rules include our social circumstances, some of which are widely shared across human groups, and some of which are more particular.

The inclusion of history in the study of morals is not new. Philosophers have long speculated about how historical factors have shaped moral values, and many leading ethicists have offered historical accounts. Prominent examples include Hobbes,

Rousseau, and Hume. The stories we find in these authors' works are in some sense fanciful, however. They offer highly speculative accounts of why values might emerge from an initial state of nature, in which moral values as we know them do not exist. No evidence for these stories is offered; they are inferred from specific views about how people act in their natural state. In the *Leviathan*, for example, Hobbes tells us that human beings are naturally selfish and violent, but relatively equal in strength, which means the state of nature is a war of all against all. Morality emerges as a solution to this unhappy form of life. Taken as an empirical hypothesis, the Hobbesian account might be investigated by analyzing our natural tendencies towards aggression (a psychological thesis), and the role of the state in reducing interpersonal conflict (a historical thesis). Some empirical evidence sits well with Hobbes. For example, Wrangham (2004) documents extreme violence in small scale societies, and Pinker (2007) argues that violence has been on a steady decline. On the other hand, the Hobbesian idea of a state of nature may be a fiction. Our species is social and has always lived with socially negotiated norms and Hobbes may also exaggerate our tendency toward violence, which is counterbalanced by a tendency to look out for members of the in-group. The claim that states have served to reduce violence is hard to reconcile with mass-scale war, imperialism, and slavery, even if recent times have seen a significant decline in mortality rates. In any case, it should be clear that empirical evidence could be brought to bear on this and other historical accounts within philosophy.

Hobbes, Hume, and Rousseau are interested in how we arrived at morality from a pre-moral position. That is an interesting question, but one which hinges on a confusion if humans form social groups by nature: There may be no pre-moral position. These approaches also pose the historical question at a high level of generality, asking about the origin of cooperation, justice, or morality in general, rather than specific norms. As such they offer little insight into why cultures have different moral values, values that can even be diametrically opposed. The philosopher most famous for addressing this question is Nietzsche, whose *On The Genealogy of Morals* (1887/2009) offers a historical conjecture to explain why Christian morals differ from values documented in ancient Rome. Nietzsche offers philological evidence for his thesis that Christians inverted the Roman value system, and he relies on basic historical facts and psychological conjecture in supposing that this inversion might have occurred because the Christians had been enslaved by the Romans. When the Christians gained power, their resentment towards their former oppressors led to a moral inversion in which Roman ideals of the good, such as flourishing, were replaced by a conception of the good that includes asceticism and guilt. Again, these are empirical claims. Is Christian morality driven by resentment? Were Christians serving as Roman slaves? There is some evidence that Nietzsche got it wrong (Prinz 2007b). The Christian revolution might have been driven by middle-class Roman converts, who were predominantly female and wanted to achieve a better life.

In any case, Nietzsche's "genealogical" approach points to an under-developed resource in studying morality. Some philosophers, most notably Michel Foucault, have offered genealogical analysis to explain contemporary values and moral

variations across time and place. But there has otherwise been little uptake of the Nietzschean approach within philosophy. Within cognitive science, the story is similar, with disproportionate resources funneled into evolutionary accounts, which do better at explaining moral universals than moral differences.

One reason for this resistance to genealogical approaches is that they may appear to be unscientific in an important sense. Science specializes in generalization, and many historical developments seem to depend on one-off events, rather than repeatable laws. For example, the specific styles of art that emerged in Europe during the course of the twentieth century reflect non-repeatable historical events and innovations by individual artists. Cubism arose, in part, because the invention of the camera freed the artist from the fetters of realism; futurism arose in part because of the rapid rise of technologies of speed; Dadaism emerged in the wake of the first world war; and so on. Some moral rules are like this, including Nietzsche's case study of Christian values. But, in many cases, the factors that influence moral values are repeatable and repeated in different historical contexts. In those cases, we can see that there is room for a cultural science of moral norms. To illustrate, let's consider some examples.

Cannibalism: Cannibalism is now reviled as the most evil activity that a human being can engage in, but it has been practiced by many societies across the globe throughout history. In one sample, more than a third of historically documented societies engaged in some form of cannibalism (Sanday 1986). Even the Christian Eucharist can be seen as a residue of a practice that was once more widespread. Given this variation, it would be nice to explain why some cultures engage in cannibalism and others do not. Harris (1977/1991) offers an explanation that appeals to three factors: size, subsistence, and resource availability. Hunter-gatherer societies who compete with neighbors over resources often end up in violent conflicts (Wrangham 2004). Victors in those conflicts end up with dead bodies and prisoners. From a cost benefit analysis, it makes sense to eat dead bodies, since they are a source of good meat and meat is hard come by. It also makes sense to kill the prisoners since it is too costly to enslave them. That means more dead bodies, which should also be consumed. Harris argues that cannibalism disappears with the rise of state scale societies. States have the power to form armies, which can collect taxes or tribute money from neighbors. States also tend to engage in trade relations, and have agriculture and domesticated animals, which minimizes resource competition and the need for hunted meats. Eating your neighbors is no longer advisable when they are trade partners and tax payers, so cannibalism tends to disappear with societal development.

Marriage: Marriage is a moralized institution. We consider some kinds of relationships acceptable and others unnatural or morally dubious. In contemporary Western societies, monogamy is morally preferred. When politician or golf stars stray, they lose votes and commercial sponsors. But, when we look beyond the West, more than 80 % of societies allow polygyny (Murdock and White 1969), so our moral attitudes toward indiscretion make us cultural outliers. Monogamy in Western Europe may result largely from a historical accident. Under the early Christian Church, there

was a sweeping set of reforms, which had the net effect of reducing the number of sexual partners by curtailing premarital sex, divorce, concubines, and polygyny. These policies reduced family size and led to increased heirlessness, which meant more money was donated to the Church, allowing it to spread its reforms farther and farther (Goody 1983). But monogamy is unusual because many common factors promote polygyny (see White and Burton 1988): Male-centered living arrangements favor male control over resources (e.g., patrilocal households), giving men opportunities to control women's lives; female contributions to subsistence, especially domestic contributions, make women a "commodity" worth collecting for men; room for territorial expansion promotes families with a large number of offspring, which again favors polygyny; warfare, which increases male fatalities and increases the female to male gender ratio promoting many-to-one marriages; warfare for plunder, which includes capture of wives can affect gender ratios and allow young men to avoid paying for brides, promoting a further increase in polygyny; restrictions on female property ownership and competition in open labor markets makes women depend on men, creating a gender asymmetry that compels women to accept plural marriages. Given widespread male dominance, it is not surprising that polygyny is the norm. But the degree of polygyny diminishes as these factors decline. For example, polygyny tends to decline with lifestyles that are less conducive to expansion, including fishing, some forms of farming, and urbanization. The Romans who were highly urbanized made monogamy the law. In settings where expansion is particularly limited, polyandry may even arise, as in traditional Tibet and Nepal. In contemporary Western culture, there is no a widespread move to allow gay marriage, which may stem from the fact that contemporary economic systems make it profitable, for the first time, to have fewer children (Werner 1979). Heterosexual couples are also marrying later, and wealthy families are having fewer offspring than the poor. Gay marriage may be part of this same syndrome.

Incest: Cultures also vary in the degree to which they permit marriage within the family. There is probably a biological predisposition to avoid some forms of incest, but only 44 % of societies have explicit incest taboos (Thornhill 1991). The presence of these taboos and the severity of the punishment correlate with social stratification, suggesting that moral sanctions against incest arise to prevent families from consolidating wealth and moving up the social ladder. There is also cultural variation in what counts as incest. The Christian Church prohibited cousin marriage up to the seventh degree, but in the Islamic world cousin marriage is encouraged. In contemporary Saudi Arabia and Pakistan over 50 % of married couples are cousins (Bittles 1990). This may have to do with the fact that power is distributed across clans in such societies, rather than centralized, as under the Christian Church. There are also conditions that favor sibling incest. This is well documented in royal families, who want to retain wealth and avoid forming obligations to other families and groups. In Ptolemaic Egypt, Greco-Roman citizens had sibling incest rates up to 30 %, presumably to avoid having to intermarry with the Egyptians whom they had conquered (Shaw 1992).

Slavery: Many societies allowed slavery, and the anti-slavery movements of the eighteenth and nineteenth centuries were virtually unprecedented historically,

especially when considering large-scale societies. Large-scale societies often placed restrictions on who could be enslaved (outgroups, rather than ingroups), but, until recently, there has been widespread consensus within such societies that slavery in some form was permissible. Small-scale societies tend not to have slaves because they cannot feed or police slaves effectively. But when state-scale societies emerge, usually through the innovation of agriculture and food storage technologies, surplus resources and power differentials arise, and labor demands increase. This makes slavery cost-effective. Goody (1980) reports that only 3 % of hunter-gatherer societies have slaves, as compared to 43 % of societies with advanced agriculture and 73 % of pastoral societies. Economic advances gave rise to new needs (e.g., a need for a large class of laborers who lack upward mobility), new opportunities for the powerful to pursue self-interested desires (e.g., obtaining fully submissive sexual partners), and the technological and human resources needed to wage war against weaker neighbors, resulting in a class of conquered captives. Given this pattern, slavery is a likely outcome of economic growth. It is surprising, then, that slavery was ultimately banned in many parts of the world, and the primary cause may have been the industrial revolution. Proponents of the anti-slavery movement in England, which helped spark reforms elsewhere, argued that an economy based on wage labor would be more profitable. In the end they were probably right. The argument was harder to sell in the United States, where slave cotton constituted up to 30 % of the U.S. economy (Davis 1984), but Northern manufacturers who had an opportunity to change the balance of power from the agricultural South had some incentive to end slavery, and that may have contributed to the American Civil War.

Torture: Judicial torture was once widely practiced in Europe. Torture was often horrifically cruel and sometimes observed by the public. It was used to extract confessions, and, less frequently, as a form of punishment. Torture is still practiced in some countries today, and Western nations occasionally debate whether certain forms of torture should be legally permitted, but there is a wide consensus now that torture is wrong. In the eighteenth century, torture came under heavy criticism and mostly disappeared (Beccaria 1764). There had been critics of torture before Beccaria, because it was often administered at the whim of lay judges, but the eighteenth century brought a more dramatic shift in thinking. Slavery was not just something that had to be carefully regulated; it came to be regarded as fundamentally wrong.

No one knows exactly what caused this shift, but several factors may be relevant. As one example, Europe endured massive losses during the 30 Years War (almost 10 % of the population died), and people were weary of violence. That, and subsequent brutal wars, helped fuel contempt for governmental use of violence, sowing the seed for an anti-torture sentiment. In the following century, there also was a shift from monarchy to more democratic forms of government. This meant that governments were, for the first time, “of the people.” When a state is led by a monarchy, it needs to establish authority, and violence is one method of doing so. When a state is led by the people, there is less need to establish authority, because the people have no difficulty granting authority to themselves. Thus, democracy may have bolstered negative attitudes towards torture.

Another variable is the perception of a foreign threat that has penetrated the sanctity of the state (Thurston 2000). European torture was often directed at people accused of heresy or witchcraft, which was regarded as a kind of supernatural invasion from within. In more recent times, torture was used during the Soviet Terror of the 1930s, under paranoid suspicion that counter-revolutionaries were secretly operating from within to undermine the state. Torture was practiced during Argentina's Dirty War, which was fuelled by fear of an internal communists threat. As part of the War on Terror, the U.S. used torture techniques against alleged foreign enemies who allegedly conspired to commit violent acts on American soil.

6.2.3 Implications

Examples of the foregoing kind are easy to multiply. They illustrate several important points. First, there is a tremendous amount of moral variation. Each value endorsed by one culture is rejected by others. This shows that morality is plastic. There are dramatic cultural differences concerning who is deemed morally worthy and in the appropriate treatment for those designated as unworthy. Thus, we must move beyond nativist and evolutionary approaches if we are to understand the beginnings of morality.

Second, moral values are essentially historical. Each has a genealogy. Thus, history is an important tool in explaining morality. Third, though many cross-cultural differences result from specific historical events, others can be explained by appeal to variables that re-appear across time and space.

For these reasons, there can be a cultural science of morals, tracing factors that can lead to the emergence and retention of some values and disappearance of others. Research on cultural evolution has moved in this direction. Cultural evolution refers to the idea that cultural items are subject to pressures similar to natural selection. Cultural items, including moral norms, vary in their degree of fitness (i.e., their likelihood to be passed on to the next generation). Fitness here can include biological fitness because some norms lead to greater reproductive success. But it can also include psychological fitness since some standards are easier to learn or more catchy. Norms that increase the power of norm-disseminators can also be said have a high degree of cultural fitness, such as norms that increased the coffers of the church. Given this broad notion of fitness, it is important to see that cultural evolution differs from biological evolution, but both forms of evolution illustrate how historical processes might be characterized by general principles, and are thus amenable to scientific inquiry.

It does not follow from this that human plasticity is open-ended. Perhaps some moral rules are easier to learn than others and some might even be impossible to sustain. Morality is no doubt constrained by our biological endowment. The emotions we have, our capacity to attribute mental states, and our care for kin all serve as building blocks that help shape the outcome of norm construction. The anti-nativist does not postulate a blank slate. But the biological constraints should not be

mistaken for a moral sense. They may constrain morality the way human visual capacities and emotions constrain the arts.

Thus, the scientific study of morality should not be limited to psychology, neuroscience, ethology, and biological evolution. It should expand to include anthropology, history, sociology, and other fields that track sources of cultural variation. A complete science of morality will work at multiple levels. Material factors will influence cultures, cultures will affect moral education, moral education will tune emotions, and emotions are implemented by circuits in the brain. Evolved human biology will contribute to this story, by shaping behavioral predispositions and the affective and cognitive faculties that allow us to internalize moral values. But this should not lead us to adopt the kind of reductionism that construes the moral faculty as a historical. To do so would be to overlook the most distinctive aspect of human psychology: how we think is affected by institutions that we create and transmit socially. Moral variation over time and the conflicts that divide the world today can be understood only if we overcome nativist biases and look at morality through a cultural lens.¹

¹I am deeply indebted to Markus Christen, Carel van Schaik, and an anonymous referee for enormously helpful comments.