Original Articles

Moral vindications

Victor Kumar

Philosophy Department, Boston University, 745 Commonwealth Ave, Room 516, Boston, MA 02215, United States

ARTICLE INFO

ABSTRACT

Psychologists and neuroscientists have recently been unearthing the unconscious processes that give rise to moral intuitions and emotions. According to skeptics like Joshua Greene, what has been found casts doubt on many of our moral beliefs. However, a new approach in moral psychology develops a learning-theoretic framework that has been successfully applied in a number of other domains. This framework suggests that model-based learning shapes intuitions and emotions. Model-based learning explains how moral thought and feeling are attuned to local material and social conditions. Philosophers can draw on these explanations, in some cases, in order to vindicate episodes of moral change. Explanations can support justifications by showing that they are not mere rationalizations. In addition, philosophical justifications are a fertile source for empirical hypotheses about the rational learning mechanisms that shape moral intuitions and emotions.

© 2017 Elsevier B.V. All rights reserved.

1. Introduction

Human beings need explanations. We're especially motivated to seek explanations for the behavior of other human beings around us, a trait that emerges early in development. You don't need to be a parent to know that children as young as two incessantly pose the question "why?" Children are even known to seek explanations for this very tendency. One three-year-old is recorded to have said to her mother: "Mommy, I always ask why. Why do I always ask why?" (Callaman & Oakes, 1992: 222; cited in Nichols, 2015: 17).

A few curious children grow up to become more serious students of human behavior. Both psychologists and philosophers wonder why we make the moral judgments and decisions that we do. In a moral context, however, "why questions" are ambiguous—between questions of explanation and questions of justification. When we ask *why* someone donated her hard-earned money to charity, we might be wondering what reasons *caused* her to do so, or we might be wondering what reasons *justify* her choice. The "reason" for a moral judgment or decision—the answer to a "why question"—may be a causal explanation or it may be a normative justification.

Explanation and justification often diverge. Imagine someone who donated to charity because she was moved to imitate a celebrity. This explains her choice but does not justify it. However, when we ask the same person why she donated, she cites a duty to help those in desperate need. This is a perfectly good rationale but does not reveal the actual cause of her choice. We have a justification, but we lack an explanation. The distinction at hand suggests a natural division of labor between psychologists and philosophers who study morality. Moral psychologists seek explanations for moral judgments and decisions; moral philosophers seek justifications.

Some interdisciplinary work on moral thought, however, blends science and philosophy, weaving empirical and normative threads into the same cloth. Sometimes an answer to a "why question" yields both an explanation and a justification—both a cause and a rationale. For example, if research suggests that heterosexuals become more accepting of gay people by empathically appreciating the harms that certain friends and family members suffer, and then generalizing their empathic insight to other gay people whom they don't know personally, the cause of this attitude change also justifies it (Kumar, 2017a). When one and the same account of moral change combines explanation with justification, it is what we shall call a "vindication" (Kumar & Campbell, 2017).

Vindications of moral change can be found in the field of moral learning. In general, learning theory in moral psychology studies the implicit and explicit learning mechanisms that shape moral judgments and decisions (along with other attitudes). Learning mechanisms typically effect stable moral change by reorganizing underlying intuitive and affective structures (Campbell & Kumar, 2012). Moreover, learning can attune moral attitudes in ways that are sensitive to morally relevant aspects of the material and social environment. Under these conditions, moral change is progressive.

E-mail address: victor.c.kumar@gmail.com

Traditionally, moral philosophers seek ethical theories—like utilitarianism—that explain why any action is right or wrong, in any and all conceivable circumstances. However, one might reasonably worry whether such "ideal theories" are genuinely knowable. A similar epistemological worry has led some political philosophers to be skeptical about theories articulating the structure of an ideally just state (Anderson, 2011; Sen, 2009). Relinquishing an ideal ethical theory, philosophers might instead focus on something more accessible: moral progress (see Buchanan & Powell, 2016). We stand a better chance knowing which remote and recent moral changes have been morally progressive (or regressive) than we do of knowing any complete and definitive moral code. Empirical work has the potential to inform non-ideal theory if it can support generalizations about how moral progress occurs, and thus help to formulate methods for achieving further moral progress (Kumar, in press-a). Furthermore, as we'll see, empirical work offers valuable constraints for non-ideal theorists pursuing philosophical justifications of moral change. Explanations can support justifications by showing that they are not mere rationalizations.

Scientists and philosophers who study moral learning can profit from one another. On the one hand, investigating why moral attitudes change can shed light on their rationale. On the other hand, exploring why moral change is rational can offer clues about the psychological mechanisms that lie behind it. The aim of this essay is to show how explanation and justification of moral change are mutually informative.

2. Debunking

My principal topic in this essay is learning mechanisms that vindicate moral change. I will begin, however, by surveying adjacent and more familiar terrain at the intersection of cognitive science and moral philosophy: arguments that attempt to debunk (rather than vindicate) moral attitudes. Later on, I will describe a new approach in moral psychology and the vindications of moral change that it promises.

According to some researchers, moral intuition is a psychological module designed by natural selection: a fast, automatic, unconscious system, relatively isolated from the rest of the mind (cf. Fodor, 1983; Sperber, 1994). Moral intuition, on a nativist view, is relatively *inflexible*: much of its contents are fixed either innately or in a critical developmental window. For example, Mikhail (2011) and Dwyer (2006) argue for the existence of an innate and universal moral grammar, triggered by social cues in early development and capable of producing only a limited variety of local moral languages.

Greene (2008, 2013, 2014) has developed a theory of moral intuition that blends nativism and empiricism. According to Greene, morality is fundamentally a biological adaptation for living in small hunter-gatherer groups that were in severe competition with one another. Moral intuition is a set of simple heuristics that are useful for coexistence within tribal groups. Intuition, Greene says, is like the "point-and-shoot automatic settings" on a digital camera: "highly efficient, but not very flexible" (Greene, 2014: 696). However, Greene also says these "automatic settings…can be acquired or modified through cultural learning…[and] individual experiences" (698).

Greene thinks that this view of moral intuition has important philosophical implications. He argues that moral heuristics often lead us astray in contemporary, large-scale, technological societies. For example, Greene's own experimental work suggests that intuition produces an emotional aversion to harm inflicted by "personal force," roughly, the use of direct, muscular force to inflict violence upon people. In the Pleistocene environment of evolutionary adaptedness, this was virtually the only way to harm someone, and so a "personal force heuristic" made good sense. But, as Greene says, it doesn't matter, ethically speaking, if you harm someone remotely rather than directly. Thus, while an emotional aversion to personal force is often useful, it is error-prone (Greene, 2014: 713). This is especially true in complex technological societies in which harm is often systemic rather than dyadic, in which death and destruction are increasingly possible through the press of a button. In sum, according to Greene, moral intuition is designed for a lost past and is therefore untrustworthy in new and relatively unfamiliar conditions (714–7). Intuition is thus similar to the psychological drives that produce hunger (697). Adapted for an old environment in which foods rich in calories were rare, these drives now lead us astray in a new environment saturated with candy and fast food.

Greene uses his theory of moral intuition to argue against *deontology*, a class of theories in moral philosophy according to which the rightness or wrongness of actions is partly independent of their consequences. Greene argues that evolutionary relics like the personal force heuristic underlie "characteristically deontological" moral intuitions. The moral theorizing of deontologist philosophers is, at base, an exercise in rationalization—an attempt to offer sophisticated post hoc justifications for intuitions that are actually based on simple heuristics. Since deontological intuitions are untrustworthy, Greene argues, so too are rationalizations of these intuitions within deontology.

At the beginning of the essay, I introduced the idea of an explanation that justifies. Greene offers the inverse: an explanation that "unjustifies." He argues that once we understand the source of certain moral intuitions, we find that the intuitions are dubious, along with any moral beliefs or theories based on them. This is not a vindication, but, rather, a "debunking" (see Nichols, 2014). Greene's explanation suggests that some of our moral beliefs are founded upon error-prone psychological processes. Consequently, he argues, we should give up these insecure moral beliefs and instead adopt those that are shielded from intuitive error. Utilitarianism, Greene thinks, is relatively free of influence from error-prone heuristics.

Greene's debunking argument against deontology has been met with very little sympathy among philosophers. Many critics deny that scientific research on moral beliefs has any potential to inform moral philosophy (see, e.g., Berker, 2009). As a whole, philosophers tend to heed the "is-ought gap." According to this Humean doctrine, there is a logical gap between empirical claims about what "is" the case and philosophical claims about what "ought" to be. We can gain all the knowledge to which we aspire about how a moral judgment *is* in fact made, but that won't tell us about how we *ought* to make it.

On the contrary, however, properly formulated debunking arguments are perfectly consistent with the is-ought gap—so long as empirical premises are complemented by normative premises (Kumar, in press-a). Greene, indeed, disavows any attempt to infer normative conclusions from only empirical premises (Greene, 2014: 711). He explicitly relies on normative premises too. Thus, Greene claims that it does not matter whether or not harm is inflicted through personal force—that this is a *morally irrelevant factor* on which to base moral judgments—and his philosophical critics, on pain of acute implausibility, cannot but accept this normative claim (Kumar & Campbell, 2012).

However, Greene's debunking argument against deontology proves to be unconvincing in light of a more complete description of the underlying psychological processes. Deontological intuitions are sensitive to a range of factors aside from whether harm is inflicted through personal force. For example, intuitions track the degree of harm inflicted, whether it was caused intentionally or only accidentally, whether it was intended as a means to an end