
Book Review

Author(s):
Tobia, Kevin P.

Publication date:
2015

Permanent link:
https://doi.org/10.3929/ethz-b-000303567

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
Philosophical Psychology 28(5), https://doi.org/10.1080/09515089.2013.871618
Joshua Greene introduces his recent *Moral Tribes* as an “ambitious book” (2013, 5). It covers a vast range of empirical and philosophical ground, including an overview of Greene’s research in the neuroscience of morality, a survey of the current state of much recent experimental moral philosophy and related areas in the social sciences, and an ethical argument for the “deep pragmatism” interpretation of utilitarianism. Greene is not only ambitious in his coverage, but also in his philosophical targets. He challenges the foundations of many historical and contemporary approaches in ethics: emotion, reason, religion, and science.

The book, intended for a general audience, will likely be received as part of a recent class of experimentally-informed works like Thaler and Sunstein’s *Nudge*, Ariely’s *Predictably Irrational*, and Kahneman’s *Thinking Fast and Slow*. But Greene’s aims are different, as his arguments and conclusions are presented as deep philosophical ones. I suspect professional philosophers may find *Tribes*, with references to current events and children’s books, distinctly unphilosophical. But under the surface of fMRI brain scans and tribal parables lurks a grand ethical argument, one whose success is worth philosophers’ attention.

*Moral Tribes* contains five parts. First, in “Moral Problems” Greene sets out two moral problems: the Tragedy of the Commons and the Tragedy of Commonsense Morality. The former of these he describes as a problem of “Me vs. Us” or “selfishness versus concern for others” (14). The moral machinery in our brains solves this problem, argues Greene, but it also creates a second type of problem. This problem, the Tragedy of Commonsense Morality, is a problem of values: “Us versus Them” (14). Evolution has endowed us with moral processing that fosters cooperation, solving “Me versus Us” problems, but this same machinery produces opposing values. While nature has given us resources to solve problems within our own moral tribe, these same resources lead to massively problematic inter-tribal conflicts.

Part two, “Morality Fast and Slow,” contains what readers familiar with Greene’s previous work might reasonably expect: utilitarianism, emotion, trolley problems, and brain-scans. Here Greene presents his dual-process theory of moral judgment. Roughly, the brain has two types of settings, one automatic and one manual. The former is efficient but inflexible, while the latter is inefficient but flexible. Greene claims, “the moral brain’s automatic settings are the moral emotions… the gut-level instincts that enable cooperation within personal relationships and small groups. Manual mode, in contrast, is a general capacity for practical reasoning” (15). The automatic settings, or moral emotions, are what evolution has provided us to solve “Me versus Us” problems. But these same settings create “Us versus Them” problems.

Parts of *Tribes* might be well described as empirically informed philosophy. Greene draws on a wide array of scientific literature to inform theories of mind and moral

---

psychology. Other parts, especially the latter half, employ an intriguing and likely more contentious philosophical methodology. Greene begins to draw on a range of empirical work to support an argument for a conclusion in normative ethics.

Part three seeks a metamorality, one that can adjudicate between differing group moralities, solving “Us versus Them” problems by establishing a “common currency” (175). Greene rejects religion, reason, and science as candidates, arguing instead for his “deep pragmatism” interpretation of utilitarianism (153). Greene’s utilitarianism (or deep pragmatism) is the moral theory that we ought to maximize happiness impartially (203). Part four contains responses to several classical challenges to utilitarianism, and part five features the application of Greene’s preferred deep pragmatist utilitarianism to issues in practical ethics.

The arguments for these normative conclusions draw from the empirical literature discussed throughout, as Greene considers when we ought to trust our automatic (emotional) moral processes and when we ought to swing into manual mode. Tribes, containing elements of philosophy of mind, moral psychology, metaethics, and normative ethics, concludes with an exercise in applied ethics including a case study in abortion (309) and a list of six “simple, practical suggestions” (350). An ambitious book, indeed.

There are several points in Greene’s argument upon which philosophers might press. One of these is Greene’s unconventional definition of morality: “Morality is a set of psychological adaptations that allow otherwise selfish individuals to reap the benefits of cooperation.” (23). This definition provides a morality that neglects some common topics in contemporary moral philosophy. Many philosophers will be averse to defining morality as a set of psychological adaptations, but even if we concede that morality is such a set, why confine it to the set of adaptations allowing selfish individuals to achieve the benefits of cooperation? Sometimes what we describe as moral is an unselfish and distinctively uncooperative action. For instance, if I find a Nazi living well in South America (and I am confident he will not do anything bad again), I may have a moral obligation to report him.

The treatment of the emotions may leave others nonplussed. Greene claims he values the emotions, but it seems their sole value is their ability to solve “Me vs. Us” problems (141). For Greene the emotions are automatic processes that often have some impressive cooperation-enhancing abilities. But that emotions are automatic processes is not enough to discard or disvalue them, and there is a large philosophical literature valuing emotions (for example) for having cognitive content, backing our moral reasons, or being markers of value. The mere automaticity of an emotional process does not exclude any of these possibilities.

Greene places a lot of weight on the existence of a common moral currency. “Without a common currency there can be no metamorality, no system for making compromises, trade-offs” (175). “Common currency” takes a number of forms, described as “a unified system for weighing values” (15), “shared values” (16), “happiness” (164), “human experience” (171), and what “can bridge the gap between Us and Them” (174). But it is possible that even without a unified system of weighing values, we can have a system for making trade-offs. One such system would be to choose whichever option has the largest number of positive values, and in the case of a tie to consider the positive values in some set order.

There are also questions to raise about the impartiality of deep pragmatism. If
“everyone’s happiness counts the same” (163), the big question is who counts as “everyone”? There is little discussion in the main text of animal experience, and a reasonable objection is that Greene neglects animal experience where deep pragmatism should, in principle, extend. But non-human animals are just one vehicle of experience excluded. Another neglected group is future persons. Thinking about happiness impartially should involve consideration of the happiness of those who will have experiences in the future, and maximizing happiness impartially will likely involve putting a great emphasis on our future effects.

If we combine this thinking with Greene’s arguments we might find deep pragmatism, ironically, best served through the emotions. If maximizing happiness impartially involves considering future persons, and if particular automatic processes (emotions) are good at solving cooperation problems, the best way to solve new inter-tribal cooperation problems (in the long term for lots of future people) may be to develop and inculcate automatic processes adept to handle these problems. As Greene notes (300), we tend to rationalize lots of our moral behaviors; this kind of flexibility is sometimes a vice of our controlled processing. Our automatic moral emotions, however, are inflexible and efficient. The best way to maximize happiness impartially may well be to bestow upon our descendents the best inflexible and efficient emotions.

At many points, Greene relies on an unconventional strategy to respond to objections to utilitarianism. He claims he is not after moral truth, but rather practical guidance. In response to Nozick’s utility monster objection and the rabbit-version of Parfit’s repugnant conclusion (a world of tons of rabbits feeling some small level of happiness could be better than our world of relatively few humans feeling great happiness if there are enough rabbits), Greene writes (in a footnote):

“I’m not claiming that utilitarianism is the moral truth. Nor do I claim that it perfectly captures and balances all of human values. My claim is simply that it provides a good common currency for resolving real-world moral disagreements. If the utility monsters and the rabbits ever arrive, demanding their utilitarian due, we may have to amend our principles. Or maybe they would have a good point, albeit one that we have a hard time appreciating” (381).

But sometimes Greene’s deep pragmatist project appears to engage with moral truth:

“utilitarianism becomes uniquely attractive once our moral thinking has been objectively improved by a scientific understanding of morality.” (189).

What could it mean to improve moral thinking objectively, if not to move closer to moral truth? It is also somewhat unclear how to interpret Greene’s discussion of debunking arguments (212, 241, 251, 283, 328) if we are not concerned with moral truth:

“If we use our scientific self-knowledge to debunk our biased intuitions, where will we end up? I believe that we’ll end up with something like utilitarianism” (328).
The review thus far has raised some critical questions, as is the common practice; unfettered praise is not for philosophy. But *Moral Tribes* deserves praise, and there is one impressive victory especially worthy of note. This is Greene’s response to the challenge of the normative insignificance of (neuro)science. Though Greene does not explicitly address this challenge, *Moral Tribes* serves as a demonstration of the normative significance of neuroscience and, more broadly, empirical science.

Some take the challenge to the normative significance of empirical science to be something like the following: no empirical result can by itself achieve normative conclusions. I suspect even Greene may not dispute this claim, and nothing in *Tribes* aims to meet this extreme version of the challenge. But there is another popular version of the challenge: no empirical results can do philosophical work in moral arguments. Throughout *Moral Tribes* Greene demonstrates successfully how empirical data about our moral brain and cognitive processing can be incorporated (with further normative premises) into arguments to achieve philosophical conclusions.

Philosophers will likely agree that *Moral Tribes* is ambitious book. It is also an impressive one, weaving empirical work from psychology, evolutionary theory, economics, and neuroscience into a profound ethical argument. Remarkably, it offers a serious philosophical argument for an intelligent general readership. The book also deserves to gain traction in philosophical circles. As a comprehensive state of affairs of contemporary moral psychology it is a valuable resource, and the book offers worthy ethical arguments – from metaethics and normative ethics to moral psychology and practical ethics. *Tribes* recognizes its ambition, and I suspect that this is just the beginning of new debates in empirically informed moral philosophy.