

# SAM HARRIS

## THE BLOG

[Neuroscience](#) | [Ethics](#) | May 30, 2011

### Morality Without “Free Will”



(Photo by

[H.koppdelaney \(http://www.flickr.com/photos/h-k-d/2554927006/in/photostream\)](http://www.flickr.com/photos/h-k-d/2554927006/in/photostream/))

Many people seem to believe that morality depends for its existence on a metaphysical quantity called “free will.” This conviction is occasionally expressed—often with great impatience, smugness, or piety—with the words, “*ought* implies *can*.” Like much else in philosophy that is too easily remembered (e.g. “you can’t get an *ought* from an *is*.”), this phrase has become an impediment to clear thinking.

In fact, the concept of free will is a non-starter, both philosophically and scientifically. There is simply no description of mental and physical causation that allows for this freedom that we habitually claim for ourselves and ascribe to others. Understanding this would alter our view of morality in some respects, but it wouldn’t destroy the distinction between right and wrong, or good and evil.

The following post has been adapted from my discussion of this topic in *The Moral Landscape* (pp. 102-110):

\*\*\*\*

We are conscious of only a tiny fraction of the information that our brains process in each moment. While we continually notice changes in our experience—in thought, mood, perception, behavior, etc.—we are utterly unaware of the neural events that produce these changes. In fact, by merely glancing at your face or listening to your tone of voice, others are often more aware of your internal states and motivations than you are. And yet most of us still feel that we are the authors of our own thoughts and actions.

The problem is that no account of causality leaves room for free will—thoughts, moods, and desires of every sort simply spring into view—and move us, or fail to move us, for reasons that are, from a subjective point of view, perfectly inscrutable. Why did I use the term “inscrutable” in the previous sentence? I must confess that I do not know. Was I free to do otherwise? What could such a claim possibly mean? Why, after all, didn’t the word “opaque” come to mind? Well, it just didn’t—and now that it vies for a place on the page, I find that I am still partial to my original choice. Am I free with respect to this preference? Am I free to feel that “opaque” is the better word, *when I just do not feel that it is the better word*? Am I free to change my mind? Of course not. It can only change *me*.

There is a distinction between voluntary and involuntary actions, of course, but it does nothing to support the common idea of free will (nor does it depend upon it). The former are associated with felt intentions (desires, goals, expectations, etc.) while the latter are not. All of the conventional distinctions we like to make between degrees of intent—from the bizarre neurological complaint of *alien hand syndrome* to the premeditated actions of a sniper—can be maintained: for they simply describe what else was arising in the mind at the time an action occurred. A voluntary action is accompanied by the felt intention to carry it out, while an involuntary action isn’t. Where our intentions themselves come from, however, and what determines their character in every instant, remains perfectly mysterious in subjective terms. Our sense of free will arises from a failure to appreciate this fact: we do not know what we will intend to do until the intention itself arises. To see this is to realize that you are not the author of your thoughts and actions in the way that people generally suppose. This insight does not make social and political freedom any less important, however. The freedom to do what one intends, and not to do otherwise, is no less valuable than it ever was.

While all of this can sound very abstract, it is important to realize that the question of free will is no mere curio of philosophy seminars. A belief in free will underwrites both the religious notion of “sin” and our enduring commitment to retributive justice. The Supreme Court has called free will a “universal and persistent” foundation for our system of law, distinct from “a deterministic view of human conduct that is inconsistent with the underlying precepts of our criminal justice system” (*United States v. Grayson*, 1978). Any scientific developments that threatened our notion of free will would seem to put the ethics of punishing people for their bad behavior in question.

The great worry is that any honest discussion of the underlying causes of human behavior seems to erode the notion of moral responsibility. If we view people as neuronal weather patterns, how can we coherently speak about morality? And if we remain committed to seeing people as *people*, some who can be reasoned with and some who cannot, it seems that we must find some notion of personal responsibility that fits the facts.

Happily, we can. What does it really mean to take responsibility for an action? For instance, yesterday I went to the market; as it turns out, I was fully clothed, did not steal anything, and did not buy anchovies. To say that I was responsible for my behavior is simply to say that what I did was sufficiently in keeping with my thoughts, intentions, beliefs, and desires to be considered an extension of them. If, on the other hand, I had found myself standing in the market naked, intent upon stealing as many tins of anchovies as I could carry, this behavior would be totally out of character; I would feel that I was not in my right mind, or that I was otherwise not responsible for my actions. Judgments of responsibility, therefore, depend upon the overall complexion of one’s mind, not on the metaphysics of mental cause and effect.

Consider the following examples of human violence:

1. A four-year-old boy was playing with his father's gun and killed a young woman. The gun had been kept loaded and unsecured in a dresser drawer.
2. A twelve-year-old boy, who had been the victim of continuous physical and emotional abuse, took his father's gun and intentionally shot and killed a young woman because she was teasing him.
3. A twenty-five-year-old man, who had been the victim of continuous abuse as a child, intentionally shot and killed his girlfriend because she left him for another man.
4. A twenty-five-year-old man, who had been raised by wonderful parents and never abused, intentionally shot and killed a young woman he had never met "just for the fun of it."
5. A twenty-five-year-old man, who had been raised by wonderful parents and never abused, intentionally shot and killed a young woman he had never met "just for the fun of it." An MRI of the man's brain revealed a tumor the size of a golf ball in his medial prefrontal cortex (a region responsible for the control of emotion and behavioral impulses).

In each case a young woman has died, and in each case her death was the result of events arising in the brain of another human being. The degree of moral outrage we feel clearly depends on the background conditions described in each case. We suspect that a four-year-old child cannot truly intend to kill someone and that the intentions of a twelve-year-old do not run as deep as those of an adult. In both cases 1 and 2, we know that the brain of the killer has not fully matured and that all the responsibilities of personhood have not yet been conferred. The history of abuse and precipitating cause in example 3 seem to mitigate the man's guilt: this was a crime of passion committed by a person who had himself suffered at the hands of others. In 4, we have no abuse, and the motive brands the perpetrator a psychopath. In 5, we appear to have the same psychopathic behavior and motive, but a brain tumor somehow changes the moral calculus entirely: given its location, it seems to divest the killer of all responsibility. How can we make sense of these gradations of moral blame when brains and their background influences are, in every case, and to exactly the same degree, the real cause of a woman's death?

It seems to me that we need not have any illusions about a causal agent living within the human mind to condemn such a mind as unethical, negligent, or even evil, and therefore liable to occasion further harm. What we condemn in another person is the *intention to do harm*—and thus any condition or circumstance (e.g., accident, mental illness, youth) that makes it unlikely that a person could harbor such an intention would mitigate guilt, without any recourse to notions of free will. Likewise, degrees of guilt could be judged, as they are now, by reference to the facts of the case: the personality of the accused, his prior offenses, his patterns of association with others, his use of intoxicants, his confessed intentions with regard to the victim, etc. If a person's actions seem to have been entirely out of character, this will influence our sense of the risk he now poses to others. If the accused appears unrepentant and anxious to kill again, we need entertain no notions of free will to consider him a danger to society.

Why is the conscious decision to do another person harm particularly blameworthy? Because consciousness is, among other things, the context in which our intentions become available to us. What we do subsequent to conscious planning tends to most fully reflect the global properties of our minds—our beliefs, desires, goals, prejudices, etc. If, after weeks of deliberation, library research, and debate with your friends, you still decide to kill the king—well, then killing the king really reflects the sort of person you are.

While viewing human beings as forces of nature does not prevent us from thinking in terms of moral responsibility, it does call the logic of retribution into question. Clearly, we need to build prisons for people who are intent upon harming others. But if we could incarcerate earthquakes and hurricanes for their crimes, we would build prisons for them as well. The men and women on death row have some combination of bad genes,

bad parents, bad ideas, and bad luck—which of these quantities, exactly, were they responsible for? No human being stands as author to his own genes or his upbringing, and yet we have every reason to believe that these factors determine his character throughout life. Our system of justice should reflect our understanding that each of us could have been dealt a very different hand in life. In fact, it seems immoral not to recognize just how much luck is involved in morality itself.

Consider what would happen if we discovered a cure for human evil. Imagine, for the sake of argument, that every relevant change in the human brain can be made cheaply, painlessly, and safely. The cure for psychopathy can be put directly into the food supply like vitamin D. Evil is now nothing more than a nutritional deficiency.

If we imagine that a cure for evil exists, we can see that our retributive impulse is ethically flawed. Consider, for instance, the prospect of *withholding* the cure for evil from a murderer as part of his punishment. Would this make any sense at all? What could it possibly mean to say that a person *deserves* to have this treatment withheld? What if the treatment had been available prior to his crime? Would he still be responsible for his actions? It seems far more likely that those who had been aware of his case would be indicted for negligence. Would it make any sense at all to deny surgery to the man in example 5 as a punishment if we knew the brain tumor was the proximate cause of his violence? Of course not. The urge for retribution, therefore, seems to depend upon our not seeing the underlying causes of human behavior.

Despite our attachment to notions of free will, most of us know that disorders of the brain can trump the best intentions of the mind. This shift in understanding represents progress toward a deeper, more consistent, and more compassionate view of our common humanity—and we should note that this is progress away from religious metaphysics. Few concepts have offered greater scope for human cruelty than the idea of an immortal soul that stands independent of all material influences, ranging from genes to economic systems. And yet one of the fears surrounding our progress in neuroscience is that this knowledge will dehumanize us.

Could thinking about the mind as the product of the physical brain diminish our compassion for one another? While it is reasonable to ask this question, it seems to me that, on balance, soul/body dualism has been the enemy of compassion. The moral stigma that still surrounds disorders of mood and cognition seems largely the result of viewing the mind as distinct from the brain. When the pancreas fails to produce insulin, there is no shame in taking synthetic insulin to compensate for its lost function. Many people do not feel the same way about regulating mood with antidepressants (for reasons that appear quite distinct from any concern about potential side effects). If this bias has diminished in recent years, it has been because of an increased appreciation of the brain as a physical organ.

However, the issue of retribution is a genuinely tricky one. In a [fascinating article \(http://www.newyorker.com/reporting/2008/04/21/080421fa\\_fact\\_diamond\)](http://www.newyorker.com/reporting/2008/04/21/080421fa_fact_diamond) in *The New Yorker*, Jared Diamond writes of the high price we often pay for leaving vengeance to the state. He compares the experience of his friend Daniel, a New Guinea highlander, who avenged the death of a paternal uncle and felt exquisite relief, to the tragic experience of his late father-in-law, who had the opportunity to kill the man who murdered his family during the Holocaust but opted instead to turn him over to the police. After spending only a year in jail, the killer was released, and Diamond's father-in-law spent the last sixty years of his life "tormented by regret and guilt." While there is much to be said against the vendetta culture of the New Guinea Highlands, it is clear that the practice of taking vengeance answers to a common psychological need.

We are deeply disposed to perceive people as the authors of their actions, to hold them responsible for the wrongs they do us, and to feel that these debts must be repaid. Often, the only compensation that seems appropriate requires that the perpetrator of a crime suffer or forfeit his life. It remains to be seen how the best system of justice would steward these impulses. Clearly, a full account of the causes of human behavior should undermine our natural response to injustice, at least to some degree. It seems doubtful, for instance, that Diamond's father-in-law would have suffered the same pangs of unrequited vengeance if his family had been trampled by an elephant or laid low by cholera. Similarly, we can expect that his regret would have been

significantly eased if he had learned that his family’s killer had lived a flawlessly moral life until a virus began ravaging his medial prefrontal cortex.

It may be that a sham form of retribution could still be moral, if it led people to behave far better than they otherwise would. Whether it is useful to emphasize the punishment of certain criminals—rather than their containment or rehabilitation—is a question for social and psychological science. But it seems clear that a desire for retribution, based upon the idea that each person is the free author of his thoughts and actions, rests on a cognitive and emotional illusion—and perpetuates a moral one.

Copyright 2011 Sam Harris