

Moral Disenchantment and the Reactive Attitudes Debate

One of the major projects of contemporary philosophy aims at reconciling the humanistic image of persons as free and morally responsible agents with the often disenchanting picture that emerges from the natural sciences. Most contemporary philosophers concede that a naturalistic worldview cannot countenance the traditional, agent-causal conception of free will that has dominated western religious, cultural and ethical thought. Still, many contemporary philosophers just assume, with surprisingly little argument and even less empirical evidence, that we can be *enchanted* naturalists: we can embrace a fully deterministic conception of human choice, reject agent-causation, and yet manage not to undermine the psychological motives and inclination central to moral life.

In this paper, I examine an emerging body of empirical research that suggests a more pessimistic conclusion. Specifically, I bring together empirical findings from psychology, experimental philosophy, and neuroscience which lend credence to what philosophers like Saul Smilansky have long warned, that illusory conceptions of free will which are lost in the merger with contemporary science nonetheless play an important psychological role in moral motivation.¹ I call this debate concerning the impact of our conceptions of agency and determinism on human moral psychology the “Reactive Attitudes Debate”, and in Part 1, I trace the dialectic from Peter Strawson to contemporary naturalists like Dan Dennett, Owen Flanagan, and Derek Pereboom.

In the second part, I present the interdisciplinary evidence that a naturalistic conception of human choice threatens features of a healthy moral psychology. Some of the most interesting and controversial evidence in favor of this thesis includes recent experiments in social psychology on the impact of determinism on moral behavior. Experiments by Vohs and Schooler,² and Baumeister, Masicampo and DeWall³ purported to show that when participants are primed with a deterministic worldview, they are more likely to cheat, less likely to be charitable and more likely to be aggressive. However, these studies have come under fire from philosophers who point to misleading conceptions of freedom and determinism employed in the priming passages. I argue that these kinds of problems are likely to pervade future empirical research in this area. The problem is two-fold. First, philosophers have not explicitly stated the central question at the heart of this dispute in a way that can focus future empirical investigations. Secondly, philosophers have not provided the experimental arms of philosophy or psychology with a clear and accessible description of the metaphysical issues needed to design an effective priming passage. In this paper, I propose to do both. I offer a methodological guideline for future research that lays out the central features of the Reactive Attitudes Debate and clarifies the content of the metaphysical and theoretical distinctions that are at issue.

This latter task is not trivial. The challenge facing future research is not only presenting the metaphysical issues in a relatively uncontroversial way, but the more complicated task of presenting these ideas in a manner that avoids leading participants into the most common and egregious theoretical confusions. For example, it is notoriously easy to conflate determinism with fatalism, as was indeed a common critique of previous experiments.⁴ The result of this confusion was to prime participants with a particularly disenchanting (and false) conception of human choice on which we play no causal role in the outcome of our choices. In general, the issues involving free will and determinism are rife with potential for confusion, in large part because these concepts are steeped in religious and cultural thought. One of the main contributions of this paper is to offer a series of correctives drawn from the collective wisdom of philosophical pedagogy in identifying and avoiding the most insidious conceptual confusions that are likely to plague future empirical research. Examples include

¹ S. Smilansky, *Free Will and Illusion*. Oxford: Oxford University Press, 2000.

² Vohs, K.D. and Schooler, J.W. ‘The Value of Believing in Free Will: Encouraging a Belief in Determinism Increases Cheating.’ *Psychological Science* 19 (2008): 49–54.

³ R.F. Baumeister, E.J. Masicampo, and C.N. DeWall. ‘Prosocial Benefits of Feeling Free: Disbelief in Free Will Increases Aggression and Reduces Helpfulness.’ *Personality and Social Psychology Bulletin* 35 (2009): 260–2.

⁴ E. Nahmias. ‘Why Willusionism Leads to Bad Results: Comments on Baumeister, Crescioni, and Alquist.’ *Neuroethics*

successfully disentangling fatalism and determinism, and properly rejecting the common sense notion of free will without implicitly undermining the very possibility of human choice.

Finally, I present the results of a new experiment utilizing the methodology I propose. In this experiment, we sought to prime participants with a naturalistic conception of human choice that both captures the deep sense in which our brains are causally constrained as well as their exceptional cognitive flexibility and capacity for norm governance. We found that participants who received this priming passage were less likely to be charitable than a control group. So even as we correct the misdirection and mistakes of previous studies, we still observe a real danger of moral disenchantment. I conclude with an interpretation of the findings, as well as suggestions for future research.