

Manuel Vargas, *Building Better Beings: A Theory of Moral Responsibility*, (Oxford: Oxford University Press, 2013), 345 pages. ISBN 978-0-19-969754-0 (hbk.). Hardback: \$55.00

Let us say that a theory of x is *revisionist* to the extent that what it tells us about how we should understand x conflicts with how most of us ordinarily do think about x . In *Building Better Beings*, Manuel Vargas develops a revisionist compatibilist theory of free will and moral responsibility. The theory is compatibilist because it implies that free will and moral responsibility are compatible with causal determinism. It is revisionist in part because, according to Vargas, there are substantial incompatibilist elements to our pre-theoretical views about these issues, which means that Vargas's theory, if true, demands non-trivial revisions to our natural way of thinking about freedom and responsibility.

The book is divided into two parts. A sizable portion of the first part deals with methodological issues unique to Vargas's revisionist approach. While there is some interesting material here, especially the criticisms of libertarian approaches to freedom and responsibility (about which I will have more to say shortly), the most innovative and important material is found in the second part. It is there that Vargas proposes and defends his unique revisionist compatibilist theory, according to which praise and blame and associated norms are justified to the extent that they build morally better beings. In what follows, I will sketch and comment on some of the main aspects of this "agency cultivation model" of moral responsibility and Vargas's overall case for it.

Assuming that there is indeed a strong incompatibilist bent to our ordinary thinking about free will and moral responsibility, Vargas suggests that we would do well

in our search for an adequate theory of these things to look for one that can accommodate the incompatibilist strands in our thought. We should therefore first look to libertarianism, the view that combines incompatibilism with the thesis that we sometimes act freely and are morally responsible for some of our behavior. But when we do, Vargas says, we come up disappointed. One of his central worries with the view has to do with a lack of empirical support. Many libertarian theories require not only that actions not be deterministically caused, but also that there be indeterminism at precise places in the causal stream leading to action. However, Vargas contends that there is no evidence that the requisite indeterminism obtains, and he is not optimistic that further discoveries will vindicate the typical libertarian requirements. The upshot, according to Vargas, is that while most extant libertarian theories may be *compatible* with a broadly scientific view of human agents, they are not naturalistically *plausible*. Other worries Vargas raises about libertarianism have to do with the role indeterminism is supposed to play in grounding freedom and responsibility. He points out that it isn't clear how adding indeterminism (even in the right places) could imbue agents with freedom and responsibility that already satisfy plausible compatibilist conditions for free agency. Moreover, even if we grant that indeterminism makes room for a sort of freedom or control that cannot be had in a deterministic universe, libertarians face the additional challenge of justifying their claim that freedom or control of that type is required for moral responsibility.

Vargas highlights what are, to my mind, some of the more serious challenges facing libertarian theories of freedom and responsibility, and libertarians would be remiss if they didn't take these challenges seriously. That said, I am more sanguine than Vargas is that the worries to which he points can be adequately addressed. This is not the place to

attempt to substantiate my optimism, however. So in lieu of such an attempt, I should like to point out that if Vargas is right about there being substantial incompatibilist elements to our natural understanding of freedom and responsibility, and if we accept his principle of philosophical conservatism (discussed below), then we should all of us be cheering libertarianism on, as it is the only position that can accommodate both our incompatibilist sympathies and our standing commitment to moral responsibility.

If the prospects for libertarian theories of freedom and responsibility are indeed as grim as Vargas makes them out to be, many will insist that we are left with only two options. We could either stick with our ordinary incompatibilist understanding of free will and moral responsibility, an option that, in the absence of an acceptable libertarian theory, would quickly lead us down the path of responsibility-skepticism, or we could opt to revise our understanding of freedom and responsibility in a compatibilist direction. Vargas prefers the latter option, and thinks you and I should too. He proposes what he calls the “principle of philosophical conservatism,” according to which “we ought to abandon our standing commitments only as a last resort” (p.73). On the basis of this principle, he argues that we should not embrace responsibility-skepticism until we can rule out the possibility that there is a viable revisionist alternative available. But, of course, Vargas believes that we have recourse to just such an alternative.

Vargas’s revisionist alternative, which he refers to as the agency cultivation model, is an intriguing and complex blend of reasons-responsive compatibilist theories, Strawsonian quality of will theory, and moral influence theories (i.e., theories according the practice of holding one another responsible can be justified on the basis of its good consequences). There are important details to the account that merit attention but cannot

be addressed here. Here, though, are the basics. According to Vargas, what makes someone a responsible agent—i.e., an agent properly subject to the norms of, say, moralized blame—is having a capacity to recognize and respond to relevant moral considerations, a capacity Vargas insists depends not only on intrinsic features of the agent, but also on features of the agent’s environment. What makes a responsible agent blameworthy for a particular bit of behavior is that the agent acted with ill will or a lack of due regard for the relevant moral considerations. And, finally, what justifies adopting and implementing these responsibility-norms and associated practices is that doing so fosters agency that is sensitive to and guided by moral considerations. In other words, the reason why responsible agents are properly or justifiably subject to blame when their behavior exhibits an objectionably quality of will is that blaming them for such behavior tends to increase the extent to which they recognize and act on moral considerations.

What to make of this agency cultivation model? Libertarians and responsibility skeptics will no doubt insist that it falls prey to some of the same worries that beset other compatibilist theories, and while Vargas attempts to address some of these worries, there is little in his discussion of these issues that promises to advance the debate between him and his incompatibilist opponents. Interestingly, though, it turns out that much of Vargas’s account remains in tact even if incompatibilism is true, and even if no one is in fact ever morally responsible for anything. For if Vargas is right that the practice of holding one another responsible can be justified on the basis of its consequences, the non-existence of moral responsibility would entail “virtually no significant deviation from the practices, attitudes, and judgment-licensing statuses that were thought to be at stake in light of the threat of responsibility nihilism” (p. 311).

Whether Vargas is right about all this depends in large part on whether our responsibility practices do in fact help build better beings. But what is the evidence that they do? Vargas suggests, not implausibly, that once norms of praise and blame are internalized, they can help motivate agents to act in accordance with the relevant moral considerations (pp. 175-177). Beyond this, however, he says little else to support his claim that responsibility norms and practices get people to pay closer attention to moral considerations. Nor is the claim obviously correct. Often enough blame is greeted not with moral improvement but with defensive, and sometimes even angry, attempts on the part of the offender to justify his or her bad behavior. At the very least, this should make us wonder how effective blaming actually is at cultivating moral concern. Moreover, even if our responsibility norms and practices do encourage a deeper sensitivity and commitment to morality, that by itself would not justify them, if it should turn out that the cost of adopting these norms and practice outweighs the benefits.

In attempting to show that praise and blame can be justified on the basis of their good effects, Vargas seeks to breath new life into the moral influence approach to responsibility, an approach thought by many to be moribund. He does an admirable job of showing how his agency cultivation model is largely immune to the sorts of worries thought to plague other versions of the approach, and while serious worries about the model remain, anyone interested in the questions of whether and how praise and blame can be justified will want read this book and think seriously about its arguments.

Justin A. Capes

East Tennessee State University

capesj@etsu.edu