

Manuel Vargas

Building Better Beings: A Theory of Moral Responsibility.

New York: Oxford University Press, 2013.

360 pages

\$61.00 (Hardcover ISBN 9780199697540)

It's rare enough when someone has something novel to say about the free will problem, but rarer still that someone stakes out a whole new position on it. Manuel Vargas' *revisionism* about free will and responsibility first began to trickle out in the literature about a decade ago and has undergone revision itself to appear in a mature form in this excellent book. The wide-ranging but meticulously delineated and closely argued position set out in this work can hardly survive recapitulation in a brief review. But as if I were describing a multi-stellar chef's culinary effort as a pedestrian restaurant reviewer, I will do my best to capture the tantalizing flavor of Vargas' position as well as providing some reflections on its preparation and *raison d'entrée* (to strain the metaphor). I will compare it to a cheaper homemade recipe of my own, though perhaps revealing more my own *déclassé* free will and responsibility tastes than anything else.

Vargas never says as much, but it seems clear to me that his approach to the free will problem is one embracing intellectual charity above all else. He ultimately rejects most familiar positions, such as incompatibilism, both hard- and libertarian, and semantically nihilist skepticism, and 'atomist' compatibilist views of various classic and psychologically hierarchical kinds, even though his own version of revisionism results in a kind of compatibilism after all. Still, Vargas throughout the book allows that all these positions have been staked out for some good reasons, and that many such reasons are traceable to folk belief about moral practices. One part of his overall project is to build a free will house of many mansions where very different positions get maximal due within broad logical boundaries, even if he pushes beyond typical terminology (10-15). Don't expect to see much belittlement of positions opposing his own in these pages, for the tone of Vargas' respectful disagreements takes his own thesis about building better beings to authorial heart.

As Vargas states in his introduction, the book has two chief divisions of labor. The first makes up Part I ('Building Blocks') which is his argument for his current version of revisionism. The second completes the book in Part II, which presents a new account of moral responsibility consistent with his revisionism.

Vargas begins by paying homage to folk intuitions about these matters in part of what he terms his 'diagnostic project' (22), drawing heavily upon the recent experimental work of Eddy Nahmias, Joshua Knobe, Shaun Nichols, and others (35-40). He uses this very mixed bag of data to bolster several points. For one, the folk are far from invariantist about compatibilism, incompatibilism, claims of responsibility, and so on, though Vargas is dubious about using even consistent subsets of data here to shore up prescriptivist views since the history of moral progress (if something like realist progress is granted) reveals that there is no reason to prefer merely received views over more reflective ones that replaced them. An important result of this work is that in several ways the folk are indeed incompatibilists about free will and responsibility, *contra* claims by many compatibilists. Part of Vargas' 'prescriptivist project' is to salvage what of these sentiments makes responsibility-sense, such as preserving roles for desert and alternative possibilities.

It is clear to Vargas, however, that libertarian accounts are not up to the job (Chapter 2). Therefore there can be no intelligibly absolute sense of desert (especially given naturalistic predilections, which is Vargas' base view) nor indeterministic alternative possibilities that are deemed necessary for responsibility. But if incompatibilism is draped defeated on the philosophical ropes, then doesn't that encourage up-and-coming nihilists of eliminativist and semantic varieties to seek the free will title? Vargas argues that when we look at comparable cases of where cherished mundane beliefs have been eventually abandoned by changing contexts of meaning and reference, *revisionism* about free will and moral responsibility is the better contender (Chapter 3).

Vargas identifies at least two types of revisionism about concepts that are possible (90). The first is *connotative* revisionism, where 'free will' would be revised according to an improved understanding of what the concept *means*, such as when we moved from believing that 'water' referred only to a liquid essential to life that filled oceans and rivers, etc. to believing that it was also a certain chemical combination of hydrogen and oxygen. The second is *denotational* revisionism, such as first using 'fish' to refer to whales and dolphins along with other similar water-dwelling animals, but then shifting our proper references to pick out only gill-bearing creatures. Vargas embraces connotational revisionism about free will based on prescriptivist arguments to the effect that, when properly normative circumstances bolster our confidence that individuals can be held morally accountable in many familiar ways, those kinds of instances constitute what we mean by 'free will.' Part II of Vargas' book engages in an exhaustive survey of what prescriptivist arguments are required to get the connotational shift moving toward this goal, though he concedes repeatedly that particular parts of a full and final account might need tweaking in light of better data and further effects of revision. Epistemically, revisionism apparently needs fallibilism as a constant companion.

Part II 'A Theory of Moral Responsibility' begins (Chapter 5) with an attempt to locate the general features of the account within the recent 'reasons-focused' literature on agency as found in the writings of (e.g.) Peter Strawson, Harry Frankfurt, Susan Wolf, Gary Watson, and John Martin Fischer. Vargas' main dissent from all these is that they are too 'atomistic' in attempting to give metaphysically or psychologically delimited accounts of a reasons-based agency that fly in the face of the more complex interplay of internalist and externalist factors of agency in a social and cultural environment (204-9). The reasons that most matter in Vargas' vision of agency have to do with those attached to actions that spring from a reflective concern about the overall effects of action (Chapter 6). One crucial job of an adequate theory of moral responsibility is therefore to inculcate in agents kinds of reasoning that consider norms of behavior in forming and choosing the proper ends of action, which in turn reinforces the very institution of moral responsibility (e.g., see p. 183-185). Thus the title of the book.

On the side of assessing moral responsibility by this theory, Vargas loosens and expands what is ordinarily seen to be an essential part of any theory of agency, namely what constitutes the power(s) of agents. Against familiar superstructures like agent-causes, reasons-responsiveness, deep selves, and so on, Vargas outlines a *circumstantialist* view of agency (Chapter 7). This posits internal abilities as certainly relevant, psychologically and physically, but not in isolation from more external social contexts and imposed social norms that may have significant influence on the individual agent, even given *ceteris paribus* assumptions about people in general or in a certain cultural milieu. Assessments of moral responsibility then are flexible across time and circumstances in part because abilities are too as functions of such. However, Vargas makes room

for practicable desert as part of blame, as well as a place for consideration of historical and manipulation effects on agents (Chapters 8 and 9 respectively).

In conclusion, I wish to close with some comments on another aspect of Vargas' book that often fills it with a deep respect for pragmatic concerns even if for the most part those concerns silently bathe his methodology by seeping in between the lines. (Vargas only directly addresses 'pragmatic' issues so-called in a few short but important passages in 'Justifying the Practice' chapter 6, p. 182; the index does not include any reference to 'pragmatism' or a cognate.) Pragmatism as a general approach usually embodies a penchant for naturalism, epistemic humility typically parsed as fallibilistic, and mostly for subscribing to the pragmatic maxim of truth as broadly a function of the effects of a concept. Measured by these benchmarks Vargas' book passes muster on all counts, with some modifications of terms. I have already alluded to strains of the first two characteristics in my review above. For the third I turn to Vargas' crucial idea of 'the work of the concept' which helps test a candidate concept of free will and moral responsibility for plausibility by assessing all of what it entails in terms of thought and practice (p. 105-109). While the work of a concept is not presented as a pragmatic *maxim* of truth, since Vargas clearly uses this principle to demarcate more from less tenable moral concepts and practices, I think it would be fair to say that it demonstrates at least the *possibility* of evincing truth about free will and moral responsibility.

Since connotative revisionism allows that reference of candidate terms may be stable ('water' refers to the water in Lake Michigan behind my campus irrespective of whether it's H₂O), but denotative revisionism requires some shift of reference (whales aren't 'fish' in a practical if cruelly demonstrable fashion – the drowning of whales can be used to separate them from permanently immersed gill-bearers), perhaps this difference signifies a preference for pragmatic explanation for the latter linguistic phenomenon, but more non-pragmatic, perhaps purely semantical or epistemic grounds for the former. Of course it is a thorny meta-issue whether pragmatic or some more pristine semantic or epistemic concerns win the day here because at the level of justifying methodology these concerns are hazy and may intertwine, and I have no particular wisdom to offer to settle such large-scale matters. However, specifically from a pragmatic perspective, issues of semantic meaning, use, and normative guidance in the practical effects or 'work' of a concept cannot be detached from one another, and thus I would favor enclosing both connotative and denotative revisionism within the boundaries of a broadly pragmatic philosophy of action theory. Pragmatism is presently my own basic and crude recipe for serving up the free will problem, and in my view Vargas masterfully cooks up a brilliantly refined version to be savored as much for its readable presentation as its very satisfying philosophical flavor. *Building Better Beings* is in any case a new fusion-style offering on the menu for enthusiasts of free will and action theory to try. I heartily recommend it.

V. Alan White

University of Wisconsin—Manitowoc