

This is the submitted version of a book review forthcoming in the journal *Teaching Philosophy* (<https://www.pdcnet.org/teachphil/Teaching-Philosophy>). Please refer to or quote from the forthcoming published version.

Four Views on Free Will: Second edition

John Martin Fischer, Robert Kane, Derk Pereboom, and Manuel Vargas

Hoboken: John Wiley & Sons, 2024; xii+260 pp.; \$49.95; ISBN: 978-1-394-16196-6

OLLE BLOMBERG

In 2020, David Shoemaker wrote regarding “the famous book *Four Views of Free Will*, published in 2007”, that it “was a model introduction to the State of the Question about free will, and it surely taught many students all the basic information they needed to know in order to situate themselves to do work in the field.” [“Responsibility: The State of The Question”, *The Southern Journal of Philosophy* 58 (2): 205–37, 206-7] I myself read the book a decade after its publication when I was a postdoctoral researcher. It had an unusual and exciting set-up: four influential philosophers of free will and moral responsibility first each had a chapter presenting and defending their own view (Kane’s libertarianism, Fischer’s semicompatibilism, Pereboom’s optimistic scepticism, and Vargas’s revisionism), and in the following four chapters, they courteously battled it out, each engaging with and criticising the others’ views. It was, and still is, an excellent advanced introduction to the field. This second edition, published seventeen years after the first, has the same authors and structure. The only structural change is an added three-page appendix—“Some Free Will Debates”—with a flowchart of various choice points and positions in the free will debate. The chapters presenting Kane’s and Vargas’s views differ substantially from the first edition. Pereboom’s chapter is also somewhat different. Fischer’s chapter is also revised but closely follows the first edition’s chapter in structure and content. The chapters with responses have, of course, all been rewritten. References and suggested further readings have been updated throughout.

Following a three-page “Some Terms and Concepts” preface, the reader is first presented with Kane’s libertarian view. According to libertarianism, we have free will and moral responsibility, and this presupposes indeterminism. Kane’s libertarianism has traditionally been classified as what is known as an event-causal view, as opposed to a libertarian view that takes agents

to exercise a kind of causal control that cannot be reduced to causal relations between events. But Kane now submits that “an adequate libertarian theory of free will must appeal to both agent-causation and event-causation, and neither is reducible to the other.” (41) The core of Kane’s view is unchanged from the first edition though: To act on our free will and be morally responsible for the action, we must be ultimately responsible for the aspects of who we are that lead us to so act. Such responsibility is made possible by indeterministic volitional processes taking place when we make torn decisions, often over prolonged periods of time. These processes eventually result in “will-setting actions” that make us responsible for who we are, which in turn enables us to be responsible for later “will-settled” actions. On Kane’s view, then, alternative possibilities are required for ultimate responsibility, but it may be that we lack alternative possibilities for most of the actions that we are responsible for.

Kane frames the problem of free will as a familiar *prima facie* dilemma. Its first horn consists of compelling reasons for thinking that free will is incompatible with determinism. The second horn is that free will seems unintelligible given indeterminism—indeterministic chance or luck seems to threaten rather than enable free will and moral responsibility. Kane’s chapter is a clear presentation of his attempt to show that the second horn is not problematic after all, at least not if it turns out that our torn decision-making processes in fact harness indeterminism in the way Kane envisions. It is worth noting though, that the problem of free will can also be framed simply as the challenge of making sense of free and responsible agency within a natural world brimming with events (deterministically or indeterministically) causing other events (for a compelling case for the superiority of this framing, see John Bishop’s *Natural Agency* [Cambridge University Press, 1989]). Our interest then lies directly “in whether free will is compatible with a broadly scientific picture of the universe”, as it is put in the preface (xi). But the dilemma framing suits libertarians, who think that determinism is especially threatening to free will. It also reflects how the problem of free will is typically conceived.

In the second chapter, John Fischer defends his “supercompatibilistic semicompatibilism” (87). According to this view, *moral responsibility* is compatible with both determinism and indeterminism (hence the ‘super-’) but *free will* may or may not be compatible with either (hence the ‘semi-’). Fischer accepts the former compatibility because of Frankfurt cases but is agnostic about free-will compatibilism because of the so-called Consequence Argument (according to which, under determinism, our actions are not up to us because they are

entailed by the remote past and the laws of nature, neither of which are up to us). What is required for being morally responsible for an action is that it is guided by an appropriately reasons-responsive mechanism that belongs to the agent. Fischer's chapter is a clear exposition of both compatibilism and semicompatibilism, and he conveys the intuitive attractiveness of these views well.

Pereboom argues for "hard incompatibilism", according to which responsible and free agency is incompatible with both determinism and indeterminism. While he agrees with Fischer that a suitably constructed Frankfurt case shows that alternative possibilities are not needed for such agency (and he offers an especially ingenious Frankfurt case to demonstrate this), he takes responsible and free agency to be directly undermined by determinism in the relevantly same way as it is undermined by manipulation by other agents (e.g. nefarious neuroscientist). While indeterministic agent-causation of the sort envisioned by Kane would make room for free will and backward-looking moral responsibility, he argues that we have no good reason to take such agent-causation to exist in our world (107-8, 202). This may seem bleak, but Pereboom argues that the practical consequences of his view are not so radical, and more beneficial than harmful. Many of our ethically significant emotions will be fitting even without free will, and a presumption of free will is not needed for justifying criminal punishment. We can continue to hold people responsible for their wrongdoings based on forward-looking considerations (e.g. good consequences) even if we cannot justify angry blame or punishment by appeal to the claim that the wrongdoer deserves such treatment simply by virtue of knowingly having done something wrong.

Finally, Manuel Vargas presents revisionism about free will. What distinguishes Vargas's revisionism from the other views in the book is his take on the *concept* of free will. According to Vargas, our folk concept of free will has significant libertarian elements. Hence, compatibilist accounts cannot capture that folk concept. And libertarian accounts are empirically implausible. However, philosophers should not assume "that the target of our theorizing must be free will as [people] have tended to conceive of it." (133) Vargas starts by asking what "the actual roles played by free will thought, talk, and practice" are (139). He then provides a revisionary naturalistic understanding of free will by sketching how a suit of our actual abilities could realize those actual roles. Vargas argues, roughly, that we have free will in virtue of having reasons-responsive guidance control of the sort that Fischer takes to be

crucial for moral responsibility. Given that this account is successful, Vargas argues that we should revise our concept of free will by discarding the libertarian elements rather than conclude that we lack free will. (Compare with how the discovery of water being H₂O rather than a primitive element led us to revise our concept of water. It did not lead us to eliminate water from our ontology.) Finally, Vargas argues that our responsibility practices, within which the suit of abilities he identifies with free will is embedded, are morally justified by forward-looking considerations (including that they have the valuable consequence of fostering that suit of abilities). But while the practices receive a forward-looking justification (they help secure the long-term benefits of social cooperation and cultivate a valuable form of agency), our responses of angry blame, guilt and punishment within those practices are justified by backward-looking considerations (such as that one violated a moral norm intentionally and with control).

After these statements of four views on free will, chapters 5-8 contain each author's responses to the others. These chapters include illuminating discussions of Pereboom's manipulation argument for incompatibilism, of what the relevant target concept of free will should be, and of whether everyday instances of holding responsible should be directly (favoured by Pereboom) or only indirectly (favoured by Vargas) justified by forward-looking considerations. Vargas's chapter also provides a closing overview of the book.

As with the first edition, the argumentative conversations between the authors make for stimulating reading. Having conflicting partisan presentations side by side is also pedagogically valuable. Furthermore, the fact that four views are given voice makes the book sufficiently broad in scope to cover the state of play in the field. This contrasts with two other recent introductions: *Just Deserts: Debating Free Will* (Polity, 2021), by Daniel Dennett (compatibilist) and Gregg Caruso (hard incompatibilist), and *Do We Have Free Will?* (Routledge, 2022), by Kane (libertarian) and Carolina Sartorio (compatibilist). [Kane's chapter in *Four Views on Free Will* is actually a reprint of his opening statement in *Do We Have Free Will?* (with the addition of one paragraph on page 41, from which I quoted above). Oddly, this is not mentioned anywhere, even though Fischer recommends the book in his "Further Reading" section.] The choice of represented views is also appropriate: Libertarianism, (semi)compatibilism and scepticism are the main rival views in the debate, and Vargas's revisionism highlights important methodological and metaphilosophical issues. All authors

write in relaxed and clear prose that makes for pleasant reading [occasionally perhaps a bit too relaxed, with references that a newcomer may find oblique (see e.g. 81, 200)].

Instructors should be aware that the book's format and set-up have some potential disadvantages when compared to a more traditional introduction [such as Helen Beebee's *Free Will: An Introduction* (Palgrave Macmillan, 2013) or Michael McKenna and Derk Pereboom's comprehensive *Free Will: A Contemporary Introduction* (Routledge, 2016)]. Central ideas and arguments are presented in several places (e.g. classical compatibilist analyses of the ability to do otherwise, Frankfurt cases) but sometimes never explained in detail. For example, while Fischer, Pereboom and Vargas all briefly explain or mention the Consequence Argument, the argument is never given a detailed presentation (in the first edition, Kane began his chapter with a presentation of the Consequence Argument, but he no longer relies on it to argue for incompatibilism). The authors' responses often repeat exposition of their own and the others' views. This redundancy feels unnecessary when reading the book cover-to-cover. But it allows a reader to engage with the response chapters with some comprehension without first digesting the initial four chapters. Having multiple authors also means that key terms are sometimes used in different ways. For instance, while the "Some Terms and Concepts" preface explains that compatibilism is neutral on whether we have free will or not, Pereboom takes compatibilism to include a commitment to us having free will (this is no serious problem, but it partly defeats the point of having such a preface). Finally, if one is striving for gender balance in course readings, the fact that all four authors are male will constrain one's choice of additional readings (one can imagine a counterfactual version of the book with, say, Helen Steward, Kadri Vihvelin and Kelly McCormick respectively championing libertarianism, compatibilism and revisionism).

To conclude, as a potential central book for a course on free will, the format and set-up of *Four Views on Free Will* have both upsides and downsides. But bearing these in mind, I can highly recommend using it. It should be suitable for advanced undergraduate or graduate students.

Olle Blomberg, Department of Philosophy, Linguistics and Theory of Science, University of Gothenburg, Box 200, SE-405 30 Gothenburg, Sweden; olle.blomberg@gu.se