

(This is a pre-print text)

Widerker, D. and McKenna, M. (eds.), *Moral Responsibility and Alternative Possibilities*. Burlington: Ashgate, 2003.

That moral responsibility (MR) requires alternative possibilities (APs) had been generally taken for granted in the philosophical tradition. More than thirty years ago, however, Harry Frankfurt argued that this longlasting assumption, which he called "the principle of alternate possibilities" (PAP), was false. According to PAP, in Frankfurt's own formulation, "a person is morally responsible for what he has done only if he could have done otherwise" (p. 17). Frankfurt argued for the falsity of PAP on the basis of some counterexamples, where we are invited to judge that an agent is morally responsible for an action of his while having no APs to it. Following him, many other alleged counterexamples to PAP have been produced. They are usually known in the literature as "Frankfurt cases" or "Frankfurt-style cases".

Frankfurt's challenge to PAP in his seminal paper raised a wide and long debate. This book contains several contributions to this debate. With the exception of Frankfurt's original paper, which is reprinted as chapter 1, the book contains only very recent articles, many of them not published before. The volume shows, then, only the last stages of the debate. This has some advantages over standard reading volumes, but philosophy teachers willing to recommend this excellent book to their students will be well advised to provide them with some wider introductory background.

The book focuses on a recent defense of PAP against Frankfurt-inspired attacks, which the editors call "the dilemma defense". The dilemma was already suggested, some years ago, by Robert Kane, with almost no echo, but it has occupied a prominent place in the discussion after Widerker's forceful presentation of it in his 1995 paper, "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities" (*Philosophical Review* 104: 247-61).

Widerker's dilemma is partly a response to J. M. Fischer's criticism of a prior defense of PAP, which, in his 1994 book The Metaphysics of Free Will (Oxford: Blackwell) he dubs, somehow disparagingly, the "flicker of freedom" strategy. Fischer's criticism is reprinted as chapter 2. The 'flicker' strategy tries to search for, and to find, some APs open to the agent in Frankfurt-style cases. This strategy gets support from the very structure of Frankfurt's original example, which features a counterfactual intervener (Black) who will ensure that the agent (Jones) does A if Jones shows some sign that he is not going to do it on his own. Since Jones shows no such sign, Black remains passive, but it seems that Jones *could* have shown that sign, and this looks like the alternative possibility that a PAP defender needs in order to resist Frankfurt's conclusion. Against this move, Fischer argued that Frankfurt cases could be designed where the sign that triggers Black's intervention is not voluntary, and not even an action. A sign of that type might be, for example, Jones's blushing at t. Jones does not blush at t and Black remains inactive. But then the 'flicker' theorist should accept the untenable view that it is in virtue of *these* alternatives (blushing or not blushing) that Jones is morally responsible for what he does. The lesson is that finding APs in Frankfurt cases is not enough to save PAP: it has to be shown that they are 'robust' enough to ground the agent's MR. This is the case, e.g., with alternative decisions and voluntary actions, but these are plainly absent in Fischer's example.

The editors agree with Fischer (rather hastily, in my opinion) that the 'flicker' strategy is hopeless, which seems to explain that, with the only exception of W. Rowe's agent-causal version of it (chapter 12), this strategy is not represented in the volume, though some other valuable versions thereof (such as Michael Otsuka's or Keith Wyma's) have recently been proposed.

Fischer's looks like a definitive objection to the 'flicker' strategy. But, according to Widerker, the Frankfurt theorist faces a dilemma: either the sign that keeps Black inactive (Jones's not blushing, in Fischer's example) is (or is associated with a condition that is) causally sufficient for Jones's decision to A, or it is not. If it is, then Jones's decision in the actual sequence is causally determined, but then the incompatibilist will

not agree that Jones is morally responsible for A-ing. Assuming causal determination in the actual sequence of a Frankfurt case begs the question against incompatibilists, who contend that causal determination is not compatible with moral responsibility. On the other hand, if the sign is not causally sufficient for Jones's decision to A, if it is only a reliable indicator of it, there is no reason to think that the decision is unavoidable. In this case, Jones may be morally responsible, but he has (robust) alternatives, and PAP is not refuted. On this basis, Widerker holds that, contrary to Frankfurt's assumption, no known Frankfurt cases constitute "IRR situations", that is, situations where there are circumstances that (1) make it impossible for the agent to avoid performing an action and (2) in no way bring it about that he performs it. On the first horn of the dilemma, (2) is false, while, on the second horn, (1) is false.

The challenge for Frankfurt theorists is to design examples that can escape the dilemma in virtue of having the following features: (a) the agent's decision and action are not causally determined; (b) the agent has no APs, or at least no robust ones, to such a decision and action; and (c) our intuition speaks clearly for the agent's MR. Several thinkers (A. Mele & D. Robb, E. Stump, D. Hunt, and D. Pereboom, among others) have attempted to construct, following different patterns, examples with these features. In chapters 3 (Widerker), 4 (Ginet), and 5 (Kane), it is argued that the new examples do not actually escape the dilemma, while chapters 7 (Mele/Robb), 8 (Stump), 9 (Hunt) and 10 (Pereboom) contain defenses and further elaborations of such examples in the light of criticisms.

Widerker's paper (ch. 3) is especially interesting. Widerker goes beyond his previous criticisms of Frankfurt-inspired attacks on PAP. Though he still seems to think that no cogent IRR situations can be construed, he now contends that, even if they could, PAP would not thereby be refuted. This contention rests on a positive defense of PAP (the "W-defense"), according to which "the belief that someone should *not* have done what he did is essential to our moral disapproval of his behavior" (p. 63). In an IRR situation, however, the agent could not have avoided doing what he did. Now, if the Frankfurt theorist still insists that the agent is morally responsible, the challenge is

for him to specify what this agent should have done instead. Given that the action is unavoidable, to expect "that the agent should not have done what he did is to expect him to have done the impossible" (ibid.). In my view, this defense of PAP is problematic, for, if IRR situations are possible, our intuition goes very strongly in favour of the agent's MR, but Widerker's defense asks us to reject that intuition.

Going deeper into the rich content of chapters 3-11 would exceed the space limits of this review. But let me make some general remarks about the dialectics involved and its consequences. Frankfurt theorists have always enjoyed an important advantage, namely that just one successful counterexample to PAP is enough to falsify this principle. However, finding an uncontroversial counterexample has proved to be much harder than it might seem. The 'flicker' and the dilemma defenses of PAP have placed strong restrictions on admissible examples. Especially hard is to combine unavoidability with indeterminism. Several options have been tried: cases with no prior signs, where counterfactual interveners have been replaced by actual blockage of unused alternative routes (Mele/Robb, Hunt), cases with internal signs that coincide with the first stages of the decision (Stump), or where the sign is only necessary, but not sufficient, for the decision (Pereboom). A common problem is that the examples tend to be quite contrived and complicated, so that they are correspondingly less likely to raise unambiguous intuitions about the agent's MR. Moreover, they tend to assume controversial theoretical commitments. The Mele/Robb's counterexample illustrates this problem. In order to meet objections, Mele and Robb have to reject both nomic subsumption and counterfactual conceptions of causality (p. 130), since these conceptions entail that, in their example, the agent's decision is deterministically caused. Besides, they have to postulate a special and problematic kind of preemption. Related difficulties affect the proposals of Stump (ch. 8), Pereboom (ch. 10), and McKenna (ch. 11).

In view of these problems, some thinkers (especiall Hunt, ch. 9) have resorted to the possibility of a theistic, omniscient God as a basis of Frankfurt cases. The beliefs of an omniscient God about our future decisions do not causally affect them, but leave us

no APs to them, either. This looks like a simple and elegant solution to the problem of ensuring unavoidability without assuming causal determinism. This is why Hunt holds that cases of this sort are "unrivalred in its capacity for eliminating all of Jones's alternatives while leaving the actual sequence unaffected" (p. 175). But they carry very heavy burdens: the very concepts of divine omniscience and foreknowledge are far from uncontroversial. Hunt admits that "divine foreknowledge might be logically incoherent" (p. 177). We may also ask what "omniscient" means. Is an omniscient being one that knows *everything* or everything *that can be known*? The latter sounds more intelligible, but then a libertarian can insist that free decisions are not among the things that can be known before they are made (cf. Katzoff, ch. 17).

An important consequence of the discussion about Frankfurt's attack on PAP has been the appearance of new theoretical positions. Besides a compatibilism that, unlike its classical forms, rejects PAP, we find new versions of incompatibilism, that deny the APs condition as well, but hold that MR is still incompatible with determinism, for the latter excludes other necessary conditions of MR, such as ultimacy of source and control. Stump (ch. 8; Linda Zagzebski is another example) represents the libertarian version of this "source incompatibilism", while Pereboom (ch. 10) defends a "hard" version: for him MR is incompatible with both determinism and indeterminism.

In chapter 12, W. Rowe defends an agent-causal version of the 'flicker' strategy. His view is criticized by Fischer (ch. 13) on the foreseeable basis that the APs Rowe identifies are not under the agent's voluntary control and so not robust enough to ground his MR.

In chapter 15, D. Copp deals with an important question, which has not been sufficiently attended to, namely the relationship between PAP and the principle that 'ought' implies 'can' (OIC). Through a detailed analysis, he joins those who hold that OIC, together with some plausible assumptions, implies PAP. I suspect that going deeper into these relationships can provide, in the near future, the basis for a strong positive defense of PAP. Copp's views have some bearing on Zimmerman's paper (ch. 16). According to Zimmerman, moral judgments can be of four different kinds:

axiological (good and evil), aretaic (virtue and vice), hipological (responsibility), and deontic (right and wrong). He contends that only deontic judgments presuppose APs. So, he sides with Frankfurt theorists in holding that hipological judgments do not require APs. However, if Copp is right that OIC, which employs the concept of moral obligation, implies PAP, it may well be (though Zimmerman does not take this possibility into account) that the necessity of APs for deontic judgments transmits itself to hipological judgments, and so to MR. Chapters 14 (Mele) and 17 (Katzoff), though still interesting, are not so central to the main debate which this book deals with.

Special mention deserves Frankfurt's contribution (ch. 18), which closes the volume. Frankfurt responds to some objections of Widerker's. I shall restrict myself to his response to one of them. He does not properly respond to Widerker's dilemma, but only to one of its horns. Widerker, in Frankfurt's words, "finds that the circumstances those examples [Frankfurt-style cases] describe do not actually make it impossible for the agent to avoid performing the action in question" (p. 339). Frankfurt's response is quite surprising (as the editors themselves acknowledge in p. 13). He holds that unavailability is not an essential component of the examples: "...What really counts is not whether an action was avoidable, but in what way it came to be that the action was performed" (p. 340). But it has been generally, and crucially, assumed that unavailability of the action for which the agent allegedly bears MR is an essential feature of Frankfurt cases. Even 'flicker' theorists have assumed this, though they try to find other APs open to the agent. Frankfurt's point seems to be that what really matters for MR for an action A is the actual sequence that led the agent to A, and not APs. But this view of MR was supposed to be a consequence of reflection on Frankfurt cases, not something that could stand on its own, for PAP is supported by very strong intuitions which Frankfurt examples were supposed to undermine. Directly rejecting these intuitions, as Frankfurt seems to be doing in this paper, substantially weakens his attack on PAP. Part of what sustains our blaming someone for an action is the assumption that he ought, and could (remember OIC), have avoided it, and this assumption about APs partly justifies our ascription of blame. Showing that the truth of this assumption is

irrelevant to the agent's MR and to our ascriptions thereof needs strong reasons, which avoidability in alleged counterexamples to PAP is unlikely to provide. All this, however, would deserve a much longer discussion.

Let me conclude by saying that no one seriously interested in moral responsibility and free will can afford to ignore this book.

Carlos J. Moya
Department of Metaphysics
University of Valencia