

## BENJAMIN MATHESON

*University of Birmingham*

**Tim Mawson. *Free Will: A Guide for the Perplexed*. Continuum, 2011.**

Tim Mawson's latest contribution to philosophy is a valiant effort to provide an introductory textbook that is accessible to those unfamiliar with academic philosophy. The book shows promise in places, but there are some issues that hinder it achieving its goals. Firstly, a major area in the contemporary free will debate does not receive as adequate a hearing as would be expected in a book of this sort. The unfortunate consequence skews the imagine of the current debate by making the position that Mawson is defending appear to be on much stronger footing than it actually seems to be. Secondly, there are many editorial decisions that do not give the feel or appearance of an introductory textbook. Finally, there is one prominent issue with the overall argument of the book.

Before discussing these issues, it will be useful to provide an overview of Mawson's position. Mawson defends a type of source incompatibilism that is very similar to Robert Kane's (1996) account, except that Mawson replaces the event-causal element present in Kane's account with a type of agent-causation. Source incompatibilism is the view that a person is morally responsible for an action or decision as long as they are the ultimate author (or source) of that action or decision. Mawson, like Kane, argues that to be the ultimate author of one's actions, these actions must have their source in a 'self-forming' action. A self-forming action requires alternative possibilities. Mawson differs from Kane in that he requires agent-causation for self-forming actions, whilst Kane requires event-causation that is cashed out in terms of quantum indeterminacies occurring in the brain. Kane's view is often plagued by objections relating to the problem of luck or arbitrariness, because it does not seem to be up to the agent, in a relevant sense, what the outcome of their decision is, despite the fact they have some alternative possibility.

Mawson starts his defence by exploring every day intuitions and assumptions about free will and moral responsibility. This part of the book is impressive in its use of appropriate examples and engaging prose. The five intuitions that Mawson uses to underpin his defence of source incompatibilism are: 1) 'Sometimes I could do something other than what I actually do'; (2) 'Sometimes I'm morally responsible for what I do'; (3) If I couldn't do other than what I actually do, then I wouldn't be

morally responsible for what I do'; (4) 'If I wasn't the ultimate author of my actions, then I wouldn't be morally responsible for them'; and (5) 'To the extent that I did not will an action under the morally salient description, I am not fully morally responsible for it.' (52) Intuitions (3) and (4) are the most important. (3) corresponds to the principle of alternative possibilities, and (4) corresponds to the source incompatibilist view that an agent must have significant influence of their actions and decisions.

A problem arises in Mawson's discussion of the principle of alternative possibilities. There has been much discussion ever since Harry Frankfurt's (1969) groundbreaking paper in which he provides an apparent counter-example to the principle of alternative possibilities (PAP). PAP is the notion that when persons are held morally responsible it is because they 'could have done otherwise' and when they are not it is because they could not have done otherwise. PAP covers many of the intuitions that Mawson uses to underpin his account, and a particular difficulty stems from Frankfurt-style cases. It is worrisome that much of the debate regarding Frankfurt cases has been omitted. Of course, in a book of this type it is not expected that every aspect of a debate is going to be covered. However, the way Mawson structures his argument relies on the strength and plausibility of this intuition, so it is problematic when an issue which directly relates to this is not given the full hearing it deserves. Frankfurt cases are set up in such a way that an agent no longer has alternative possibilities yet intuitively is still morally responsible. Mawson's reply is that the agent would still have some alternative, i.e. they could still have done otherwise, because it must be that the agent will show some sign of doing otherwise in order to allow the intervener to force the agent to do as they will. Mawson counters John Fischer's (1994) claim that such a 'flicker of freedom' would not be robust enough an alternative possibility to ground moral responsibility by saying that even this alternative is enough to change the outcome from one of responsibility to non-responsibility for the agent in the example. Mawson, however, does not consider Fischer's other point that it is what happens in the *actual sequence* of events that relates to moral responsibility. In the alternative scenario the agent would have lacked moral responsibility, but in the actual scenario they would still have moral responsibility whilst lacking robust alternative possibilities. Even in a textbook style book such as this there should really be a greater discussion of what is one of the two liveliest debates in contemporary free will literature (the other being the debate over manipulation arguments against compatibilism).

PAP leads to another issue with Mawson's overall argument. As mentioned earlier, much of his defence of source incompatibilism relies on its intuitive plausibility. Source incompatibilists are not required to defend PAP because they admit that agents can be morally responsible for actions that are determined to a certain extent, as long as those actions have an indeterministic source. Mawson describes the situation as holding the person morally responsible for the earlier action, i.e. the self-forming action, rather than the actions that result from the self-forming action. Although this might be true in some cases, such as Mawson's case of the drunk student (we hold the student responsible for getting drunk rather than the acts committed whilst drunk), it is not clear that this is how the social practice of moral responsibility works. It seems that normally when a person is held morally responsible, it is for the actions in question, rather than the formation of their character. Some might even argue that persons have little to no control over the formation of their characters. Even so, describing the situation as holding the person morally responsible for the actions that stem from the self-forming action seems like an equally good explanation of what is occurring in this sort of explanation. The worry that arises with this part of the argument is that PAP says that an agent is morally responsible if they 'could have done otherwise'. Mawson, following Kane (1996), sees PAP being entailed in the ultimate authorship, i.e. ultimate authorship requires alternative possibilities at moments in a person's life that will contribute towards shaping their character in order to be 'self-forming' actions. The worry arises in what appears to be sort of 'sleight of hand' as support for PAP is transferred onto the source incompatibilist view: a morally responsible agent must be the ultimate author of his or her actions. Although both positions are similar, they are not the same. Ultimate authorship requires that agents 'could have done otherwise at key moments in life', whilst PAP only requires that agents 'could have done otherwise'. A closer look at PAP is likely to reveal that the strength of PAP relies on the idea that a morally responsible agent 'could have done other for *every* action and decision'. There is no room to defend this view here, but it is worth noting that many other source incompatibilists do not require PAP for their views. For example, Derk Pereboom (2001) actually argues in favour of Frankfurt-style counter-examples against PAP. Although it is matter for further investigation whether or not intuitive support for ultimate authorship is equal to that of PAP, at the very least this issue needs to be confronted.

Much of the rest of the book covers standard ground in defence of source incompatibilism. As 'self-forming' actions play an important role in his account, Mawson spends a lot of time defending their coherence, whilst modifying Kane's concept by substituting the event-causal process that Kane is an advocate of with a form of agent-causation. On Mawson's view, agent-causation can provide the required alternative possibilities that make self-forming actions self-forming. Mawson strays from the standard defence of source incompatibilism when he invokes 'souls' as a way of explaining how agent-causation might be possible. This is an interesting way to handle the difficult the issue of agent-causation, if empirically questionable. Souls would provide a way of explaining this, but many would also complain that this sounds like simply explaining one mystery with another mystery.

The other issues are weighted towards what appear to be editorial decisions. One initial problem is the length of the chapters. Many of those unfamiliar with academic philosophy, and even many who are familiar with it, are going to be anxious at the length of each chapter. This worry is offset initially as there many subsections dividing the separate issues discussed under the topic of each chapter. Unfortunately, these subsections are neither numbered nor named. This creates great difficult when navigating through the text, especially when an issue has been raised in a specific chapter and the reader wishes to return to the section in which it was discussed. This is likely to provide difficult for those who are inexperienced in reading philosophy and who have not been conditioned into writing notes whilst they read.

A final worry regards further reading. It has become commonplace in textbooks of this sort to include sections which detail the works of others that the reader is free to explore if they are interested in a particular debate or argument. Sadly, this book lacks clearly set out sections that discuss further reading. Mawson does helpfully provide some examples of further reading, but these are in the endnotes, and many who are new to reading academic texts may not bother checking the endnotes. Providing a brief discussion on further reading allows the reader to avoid what can initially seem a daunting task of finding other areas of literature that discuss what they are interested in. Books and articles mentioned in further reading also have the benefit of being recommended by the author of the book they are currently reading.

Mawson does well to provide an introductory text that is accessible to those unfamiliar with academic philosophy. Mawson's array of

examples is novel and is described in such a way that it avoids the difficulty that might occur to those unfamiliar with many of the strange thought experiments that philosophers employ. However, the lack of further discussion on the issue of Frankfurt-style counter-examples is worrisome. This could have been countered to some degree by providing a section that detailed further reading, particularly one which pointed out the philosophers in favour of both positions. It is unfortunate that much of the intuitive support that Mawson gathered stemmed from PAP, because this required a much fuller discussion of Frankfurt-style counter-examples. These issues could have been avoided, but sadly were not, and this results in the book failing to achieve its full potential.