



ORDER: GOD'S, MAN'S AND NATURE'S

Free Will: A Guide for the Perplexed T.J. Mawson. New York: Continuum Press, 2011.

Reviewed by Michael Tiboris, USCD

T.J. Mawson's Free Will: A Guide for the Perplexed is, at once, an introduction to the free will problem and an argument for a particular version of agent causal libertarianism. Mawson is a clever writer, and the short book is filled with interesting examples, including some novel approaches to the traditional problems. The central material of the book is divided into four long chapters. The first identifies a series of commonsense intuitions about freedom and responsibility, which Mawson claims, make us natural libertarians. The second and third chapters take up the traditional conflicts in the free will literature. These are incompatibilism, the claim that the kind of free will necessary for moral responsibility is incompatible with causal determinism, and indeterminism, the claim that the collective state of the natural world and its laws do not leave only one physically possible option for action. The fourth chapter addresses "ultimate authorship," or the claim that a person's freedom and responsibility are importantly dependent on being able to understand his actions as having their ultimate origin in his own agency.

There is plenty of precedent for books like this, which are intended to bring people into the topic for the first time through making a positive argument for a preferred position. Michael Williams' excellent *Problems of Knowledge* is an example of this in epistemology, and Russ Schafer-Landau's *What Ever Happened to Good and Evil?* is an example from ethics.¹ When compared to books like these, *Free Will: A Guide for the Perplexed*, feels a bit unbalanced. Like these books, Mawson's serves these dual purposes, so it makes sense to evaluate it along each of these dimensions, which I will try to do while giving a critical summary of the text, below. Ultimately, as an introductory text, I think the book is filled with engaging examples and is often well argued. That said, the goals of the text sometimes seem to be in tension more than in

¹ Williams, Michael. *Problems of Knowledge*. Oxford University Press, 2001; Schafer-Landau, Russ. *What Ever Happened to Good and Evil?* Oxford University Press, 2003.

other, similarly styled, introductions . Each of the tasks draws a bit of energy from the other. At points, the discussion of traditional views feels incomplete and a bit unevenly focused on the strains of argument necessary for Mawson to make his case for libertarianism. Then, when we get to his preferred view, the details of the defense are abbreviated and leave much unanswered. This kind of balancing is a difficult task, however, and the book covers a lot of ground efficiently and with some clever new twists.

The book's main text begins by identifying several commonsense intuitions about our ability to act and decide. These include the belief that one could have done something other than what she actually did and, that if she couldn't then she wouldn't be morally responsible for her actions. Together these assumptions are the basic claims of libertarianism. Mawson also adds some other commonsense ideas intended to support this view, including the intuition that one must be the ultimate author of her actions if she's to be responsible, and that one is only responsible for some subset of the possible ways to describe an action she performs. Since these intuitions add up to libertarianism, Mawson suggests we are natural libertarians. The next two chapters consider two claims that non-libertarians might challenge our intuitive belief in libertarianism. The first is the belief in the incompatibility of determinism and free will. The second is the claim that our actions are undetermined. If libertarianism survives these attacks, Mawson suggests, it will remain "innocent."

Throughout the book, Mawson returns to the claim that libertarianism is intuitively plausible and, therefore, has some sort of claim to be accepted as correct, unless detractors can show that assumptions made by libertarians are false or raise objections that have to be answered (54, 61-64, 112,114,143, 169). Pedagogically, starting from intuitions is a valuable approach, and plays well for the free will debate. Additionally, this sets out a nice narrative arc for the book, beginning with a single view and then setting out to analyze the topic generally through testing its strengths in comparison with other views. The notion that we are "natural" libertarians has been pressed recently by experimental philosophers.

The notion that we're natural libertarians adds some complications as well. In particular, I had a hard time squaring the role of the "innocent until proven guilty" claim with other arguments Mawson seems to think are necessary to prove the case for incompatibilism. By itself, the fact that our commonsense leans libertarianism obviously doesn't prove the truth of libertarianism as a metaphysical claim. Mawson agrees with this, but he does think it establishes some kind of an intuitive bias in favor of it that is relevant to the dialectical burden of proof. Mawson uses the legal metaphor of "innocent until proven guilty" (61-64). The legal metaphor works fine at a sufficient level of abstraction. But if our intuitions about libertarianism

actually provide a premise in an argument for the truth of libertarianism, then the sense in which libertarianism is “innocent” has to be translated into claims about the justification we have for believing its constituent claims. Mawson attempts this at the beginning of chapter three through a discussion of the epistemic basicity of our belief in incompatibilism.

Belief in incompatibilism is basic -- we believe it without argument (58). But whether we’re justified in believing in incompatibilism depends on whether it’s true that freedom is incompatible with determinism (58-59). Therefore, we must look for positive arguments for the truth of indeterminism. This argument, it seems to me, creates confusion about the role of the innocent-until-proven-guilty argument throughout the text. Sticking closer to the legal metaphor, the thought would be that the basicity of our belief in incompatibilism, added to our intuitions about its truth, make a presumptive case for the view. If there are objections to the view, and they can be defeated, this will not prove the case in favor of incompatibilism, but the balance of the argument will still tilt toward libertarianism (63). On the other hand, the point of the epistemic basicity section seems to be that basic belief in incompatibilism is only ultimately justified by the *truth* of incompatibilism. As a result, we need positive arguments (or evidence of some other kind) in favor of its truth, not simply rejections of objections (60). But if this is correct, it renders the legal metaphor otiose. Even if we discover that the compatibilist arguments are no good, we’ll have to give arguments for believing in indeterminism. The fact that we treat incompatibilism as basic doesn’t mean that we’re justified in doing so, even if, as Mawson suggests, no really convincing arguments can be given here, since belief in incompatibilism is just too basic. Nevertheless, positive arguments seems to be what the conclusion of the section requires, and if they can’t be given, then we should embrace agnosticism rather than fall back on libertarianism.

I think Mawson’s intent here is to identify the kinds of objections and arguments that have to be made in order to turn the commonsense observations of the first chapter into a defensible metaphysical theory. Incompatibilists can’t simply rely on their intuitions because their basicness doesn’t justify belief in them. But the intuitions aren’t worthless, and collectively they produce a picture of the world that is plausible and functional. All of this could have been said far more directly and simply, however, which should be a goal for an introductory text such as this. Using two arguments, the legal metaphor and the basicity argument, so close together when they pull in different directions makes it difficult to see what’s at stake in the book’s narrative.

Another problem with claiming we’re natural libertarians lies in the selection of the intuitions themselves. As Mawson himself notes, some people might find compatibilism more

intuitively plausible. Even more likely than this, I think, many of us have intuitions that are just as commonsense as any of the ones Mawson highlights, and which show free will to be a *problem*: an inquiry that presents itself as deeply in conflict at its most basic levels. For instance, I think many people have the intuition that the world is composed of events are necessitated by their causes, and that human choices are events in the causal order. There's no reason to think that this makes us intuitive or "natural" determinists, and clearly we can offer arguments to show the intuitions are wrong. But I think you could get these intuitions out of people with no more difficulty than the intuition that "to the extent that I did not will an action under the morally salient description, I am not fully morally responsible for it" (52). If we add these intuitions to the mix, our view of free will looks like a problem. This wouldn't be a fair criticism if Mawson just used the intuitions as a starting point for discussion, but the first premise of the argument for libertarianism is supposed to be that it is presumed "innocent" by being a plausible story about our intuitions. Later on, Mawson argues that the evidence for causal determinism is inconclusive and falls back on the innocence of libertarianism. But non-libertarians would likely find this dissatisfying if they had the sense that this innocence really couldn't be established in the first place.

Mawson follows many contemporary authors in taking the relevant incompatibility to be between determinism and responsibility, not determinism and free will. If our intuitions are libertarian, then we are committed to this incompatibility. The bulk of the third chapter is concerned with the classical compatibilists' use of conditional analysis as a response to the consequence argument for incompatibilism. Conditional analysis is the idea that free will might be compatible with determinism if we could just get the right sense of "could have done otherwise" described in terms of relevant counterfactuals or nearby possible worlds. This is a traditional debate in the classic literature concerning free will and determinism, and it deserves to be treated in an introductory text.

Mawson does a fine job of condensing this difficult literature in an accessible way. At first, I was a little skeptical that conditional analysis deserved the level of detail Mawson gives it. While conditional analysis is an important part of the historical dialectic about free will, it hasn't played a large role in most contemporary discussions of moral responsibility, which have largely eschewed counterfactual analysis for discussion of the possibility of moral responsibility in the actual sequence of events. Modern theories tend to take the general failure of this approach as a reason to move toward the sense of (in)compatibilism Mawson is using. This topic is also passed over in some treatments of free will, I think, because it smells of dialectical quagmire. But Mawson's retelling of the consequence argument and the compatibilist response is clear

and people new to the topic will find it helpful. The traditional compatibilist response here is that we can do otherwise in the sense that, had we desired to (and perhaps made an effort to) do something else, then we would have succeeded. This counterfactual analysis picks out the important difference between actions a person could have done (if they had wanted to) and those that they couldn't. Had I wanted to, I could have put the book down, but whether I had wanted to or not I couldn't have flown to the moon by flapping my arms.

Mawson's response on behalf of the incompatibilist is a little harder to assess. The general form of the argument resembles some given by Chisholm.² If the compatibilists' conditional analysis is correct, then it will allow cases of people who could have done otherwise in the conditional sense, but couldn't possibly have done other than they did. This will happen in cases where a person is *incapable of wanting* to do other than they do. In such a case, it will be true of them that they could have done otherwise (if they had wanted to), even though they were incapable of wanting to, and so therefore couldn't actually have done otherwise.

Mawson makes a novel argument based on this general tactic (80-82). He suggests that if the compatibilists' argument is correct, then it will commit them to allowing that we could in fact change the past and the laws of nature. The thought is that, like the psychological incapacity to want to do otherwise, the past and the laws of nature make it impossible for us to want anything other than what we do. If it's true that had we wanted to, we could have done otherwise, then it must be true that we can change the past or the laws of nature (since that's what is required to want to do otherwise). So the classical compatibilist is stuck either accepting the implausible claim that we can change the past or the laws of nature, or that we don't live in that kind of (deterministic) universe. If I've got him right, here, Mawson's argument is formally similar to those offered by Chisholm, but usefully formulated to reference the consequence argument explicitly. These kinds of arguments ultimately proved devastating to classical compatibilists, as Mawson's examples show.

The discussion of more modern compatibilist views is too short. P.F. Strawson's *Freedom and Resentment*, and the compatibilist developments that followed are very important, and raise a lot of issues that Mawson himself seems to be concerned with, most notably authorship. Most contemporary compatibilists will grant with Mawson that the ultimate authorship of our actions is necessary for moral responsibility. What they deny is that moral responsibility and ultimate authorship requires the ability to do otherwise. Mawson's use of Dennett's now classic example of Martin Luther's declaration makes this point (96-100). He

² Chisholm, Roderick. "Human Freedom and the Self." Reprinted in Watson, Gary *Free Will*. New York: Oxford University Press, 2005: 34-37.

suggests in response that, due to responsibility tracing-like considerations, it's difficult for compatibilists to play up the importance of ultimate authorship without also addressing the importance of doing otherwise. Compatibilists, as I'll suggest below, are not without resources in responding to this. Many of them think of authorship in ways that do not require this kind of reliance on earlier libertarian-free choices.

A brief discussion of Frankfurt cases follows this. Frankfurt described cases where a person deciding between two actions will be (unbeknownst to him) forced to do one of them (A) by an external intervener should he form the intention to do the other. But, as it happens, the person decides to do A for his own reasons. In such cases, Frankfurt concludes, the person couldn't do other than A, but we still think he's morally responsible. This type of example is importantly consistent with, and may even highlight, the ways in which authorship of the action is important independently of the possibility of doing otherwise. Mawson's main objection to the Frankfurt cases is a version of what is sometimes called the Kane/Widerker objection.

[If] it is obvious that there is nothing at all the person concerned could have done in even the tiniest detail different from whatever it is he or she ended up doing, then we ineluctably decline to hold the person morally responsible for doing what she did. ... [and] we are supposing that there is something which he or she could have done other than whatever it is he or she actually did even if that other thing is simply a small 'wavering' ...

This is a compelling challenge to these kinds of arguments. Because the Frankfurt cases have been so influential, it would have been nice to hear more about the potential responses to this thought. Such arguments have been given in other introductory-type texts like *Four Views on Free Will*.³ The Frankfurt-inspired compatibilism cuts close to things that Mawson wants to say, for example that an action's source matters even within the actual sequence of events. Other influential compatibilist authors, John Martin Fischer, for example, have argued that responsibility depends on an agent's capacity to respond to patterns of reasons. This coheres with Mawson's belief that responsibility is dependent in some way on critical rationality (145-146). Mawson also notes that "hierarchical desire" views like Frankfurt's might be committed to a problematic kind of conditional analysis. This is an interesting suggestion and it could have been pursued further. All of this is by way of saying that the contemporary debate about compatibilism is a richer one than Mawson's treatment suggests. I think those new to the topic would probably find the contemporary developments more attractive than conditional analysis which, even by most compatibilist lights, is a failed project.

There is a theistic thread that runs through this book that it would not have been out of place to develop a bit more. There is reason to be concerned that the traditional theistic god is

³ pp. 59-61 (for Fischer's response) and pp. 89-92 (for Pereboom's).

incompatible with an indeterministic universe. Mawson raises the worry that God's essential praiseworthiness might be problematic for indeterminism. If determined actions can't be praiseworthy any more than they can be blameworthy, then doesn't the fact that God couldn't do other than the best things mean that he isn't really praiseworthy after all? These types of questions are great introduction to free will material. There are other places, even in this chapter, where we might consider God's place in the order of a deterministic (or indeterministic) universe. The Kane/Widerker objection, for example, presents a problem for God's omniscience. If the facts of the world up to any given point don't totally determine what will happen next, how can God have foreknowledge? These questions are old as the trees, and it would be nice to hear a bit about how authors more contemporary than Augustine think about these things. Mawson's solution to the problem of God's necessary praiseworthiness is fairly abbreviated, and ends up sounding scholastic and dependent on definitions of the divine drawn from tradition. I'm sure that he has much more compelling things to say about this, but I can see reasons for wanting to keep the conclusion of the book independent of a particular religious outlook.

The subject of the fourth chapter is "indeterminism." Here the question is whether there is any reason to believe that the world is deterministic. For the most part, these arguments have come in the form of abstract reasoning about the relationship between causal events. Reasoning of this kind, for example, is the way that Kant generates his antinomies of freedom. Mawson, perhaps rightly, is more concerned with whether there is (or could be) any empirical evidence of determination. To this end he considers the idea of deterministic natural laws, the meaning of discoveries in quantum mechanics, and the psychological or neuroscientific causes implied by the famous Libet experiments. The focus on the empirical is nice and introductory readers will find it reassuring to see that the free will literature can take science seriously. The conclusion, however, is that our knowledge of the world doesn't decisively break either toward confirming or denying indeterminism.

The chapter ends with a strange but intriguing argument meant to push the dialectical balance back toward libertarianism. Assuming that determinism is incompatible with responsibility, which Mawson takes the third chapter to have satisfactorily proved, then the fact that we're responsible implies that we can't be determined. This is a strange argument at least because the conclusion seems so much stronger than the premises. The premises are facts about ethical practice and the relationship between concepts, the conclusion is about the metaphysical make-up of the world. This, I'm sure, Mawson would reject. Like many libertarians, he thinks of moral responsibility as a deep property of agents. Of course, this is

controversial, and I think many non-libertarians rightly have serious misgivings about thinking of responsibility in this way. Moral responsibility, as an ethical status or practice, may have necessary conditions that are metaphysical (e.g. control or libertarian free will) without itself being a metaphysical property of persons. The concept of moral responsibility cannot obviously be used in the way Mawson intends. The premise that needs to be defended here, then, is that we are morally responsible in the deep desert sense. As Mawson notes, we can't defend this just by seeing that we are disposed to hold people morally responsible in certain cases. Skeptics will point out that it might be reasonable to think a criminal deserves to be punished, just as it may have been reasonable for 17th century villagers to believe in witches. No arguments about our practices will help here if the relevant notion of moral responsibility is the deep metaphysical one.

The final chapter concerns ultimate authorship, the idea that an agent's actions are only free if they find their source in the agent and are not caused by events external to her. This is a theme that runs through much of the contemporary and ancient literature on free will. The ultimate authorship condition is meant to be a jointly necessary condition for responsibility along with determinism. Even if the world is indeterministic, this wouldn't be enough for responsibility. An indeterministic universe could simply include lots of randomness, which is as much a challenge to moral assessment as a totally deterministic universe. Mawson largely uses the chapter as a way to introduce agent causationism as a plausible theoretical framework, and doesn't directly address many objections. This is acceptable for the purposes of an introductory text, since it would largely return the discussion to topics of previous chapters. Though, given the level of detail with which he treats arguments in those chapters, this chapter seems comparatively less rigorous. Nevertheless, I found it to be the most engaging.

Ultimate authorship, as Mawson understands it, distinguishes our actions from mere events that happen. If the world is indeterministic, then its collective state at any point in time does not totally determine what will happen next. This leaves an opening for an agent to make a contribution to what does ultimately happen, some kind of causal "oomph" which makes the difference. This "oomph" is not a mere event, it is contributed by the agent *qua* agent. If the "oomph" that makes the difference comes from the agent, and not forces external to her, we may say that the agent caused the event to happen. If it's an event for which one can be assessed as morally responsible, then her having contributed the "oomph" in this way is what makes her eligible for assessments of praise or blame. Mawson considers some intuitive evidence for thinking that we do add "oomph," whatever it is, to our actions, and concludes with a brief aside about what the non-reducible top-down causal source of this "oomph" could be.

He notes that this view is compatible with the traditional notion of a soul, or maybe more naturalistic emergent causally-efficacious substances. As I said earlier, these theistic considerations seem to be in the background throughout the text, and it would have been nice to see them considered a little more openly.

To Mawson's credit, the chapter serves as a very nice high-altitude introduction into the agent causationist worldview. He carefully explains the idea that this kind of authorship is (and must be) compatible with a certain amount of external causation. This is because we are certainly influenced in how we apply our agency by our characters and the situation we find ourselves in. We're not totally free to choose to do *anything whatever*. As Hume noted, in a way our decisions would be incomprehensible and certainly not morally attributable if they aren't determined by features of our psychology in some way. So, our decisions must be, as Mawson puts it, causally inclined to go one way rather than another, but not causally determined to do so. In that moment of indetermination, agents can exercise a capacity to add causal "oomph" to the action. In order to keep this from collapsing into a probabilistic version event causationism, the source of the "oomph" must be a person, an agent, not an event.

The essence of ultimate authorship is that, where it is present, I in fact cause (even if other things would have caused had I not) whatever it is I do, ... [and] that nothing outside of me causally necessitated that 'I' do whatever it is I do, but not that nothing outside of me causally inclined me to do whatever it is I do ... It just has to be the case that in fact one of its causes was me. If I add my own causal oomph to the world in this way, the resulting movements of my body and the outcomes that they bring about are actions that I am performing using my body, rather than merely events that my body is undergoing. (154)

If indeterminism and ultimate authorship are two necessary conditions for freedom and responsibility, then this account emphasizes the relationship between these conditions. Indeterminism is essential for ultimate authorship, in other words.

Agent causationism is one of those theories that looks great at a high level of abstraction, but when analyzed more closely feels mired in details and theoretical epicycles that make it far less attractive. There are obvious questions here about what it means to contribute "oomph." The view also seems to demand an explanation of what it means for causes to be "outside" or "inside" an agent. But even so, the idea that ultimate authorship of some kind is required for responsibility is appealing. Compatibilists might respond to this by trying to expose the idea of an agent cause, and causal "oomph," as reducible to event causes, thereby planting the agent back into the stream of causally determined events. Mawson briefly considers this argument, but his rejection of it largely rests on assertions defended by footnotes to other texts. More importantly, through accepting that ultimate authorship is a requirement for free action and

responsibility and deny that being an ultimate author requires indeterminism. In other words, they may deny the essential link between the two conditions Mawson gives for freedom. In doing this they might well be able to capture a lot of what Mawson thinks is important for responsible agency. For example, they will agree about the relationship between responsibility and rational reflection, and locating the source of our actions in a subset of an agent's psychological properties. It seems to me that a large portion of the contemporary literature on compatibilism is attempting to do something like this by explaining the ownership conditions for responsibility. Our actions can be the result of competent judgment, reflection, solid states of character developed through longtime habits. All of this is possible within the actual sequence. If these structural, or maybe better, *hexis* views of freedom make sense, then they will agree with Mawson about the importance of authorship, though not about the importance of indeterminism.

In the course of offering evidence for the idea that we add agent causal "oomph" to our actions, he notes that it is our everyday experience that we do this. Agent causal "oomph" explains how we can make decisions when nothing is on the line and we have no strong preference one way or the other. Determinists might respond that we could surely come up with an event causal explanation for individual choices, even as we don't experience them this way--agent causation is an illusion. "But we cannot in general," Mawson writes, "take the fact that an experience of a certain sort can -- in, not, contrived circumstances -- be illusory as a reason to suppose that it is generally illusory" (159). In the absence of evidence for this we "have no reason at all to suppose that this sceptical view *is right*" (186n22). The threat of responsibility-incompatible determination, however, isn't limited to event causal determinism. Psychologists have unearthed disturbing evidence of authorship-undermining conditions. Intuitively insignificant situational conditions seem to have a large impact on behavior. People hyperbolically discount in favor of smaller rewards at a sooner time even when they rationally know that waiting for the larger reward would be better. We confabulate and tell ourselves stories about the reasons we considered when deciding when we did nothing of the kind. These findings all challenge our perception that we are self-authoring, and as such present a problem for all parties to the debate.

As a final note, I wish to say a few things about the book's presentation. The book is filled with thoughtful examples. Some are familiar, but even these are often put to novel uses. For example, the case of Martin Luther's declaration that he could "do no other" is used to discuss, among other things, the issue of responsibility "tracing." While the examples are quite vivid, they are just as often overly long and do not always move the discussion forward. There

is, for instance, a long retelling of the sheriff who can maximize utility by convicting an innocent man; a classic objection to utilitarianism. The example occurs in a discussion of what it means to be morally responsible, and Mawson has added some nice little twists to it. But its placement and the discussion that follows encourages a slide between moral responsibility and the justifiability of punishment which complicates the issue, rather than clarifying it. Mawson has a propensity for tangent, as well. Sometimes this takes us into deeply interesting places which paint old issues in a new light, but other times it disrupts the argument of the text in ways that don't seem terribly necessary and will be difficult for novices to follow.

Mawson's *Free Will: A Guide for the Perplexed*, struggles a bit with the dual roles of introduction and self-standing text. It covers a large range of difficult and wide-ranging material well, with colorful and rich examples. The truly uninitiated will find the book very challenging. The commonsense notions of the first chapter quickly give way to a detailed and unflinchingly philosophical analysis of some of the key problems in the free will debate. To my mind, the book does not consider new developments in compatibilism with enough detail. The task of making an argument for agent causal libertarianism seems to draw attention away from giving a comprehensive introduction to the topic, and the task of writing an introduction prevents delving into the kind of detail one would need to truly defend an agent causal view. Mawson has produced an engaging book that I think advanced undergraduates with some background in philosophy will find quite useful.

Michael Tiboris

University of California, San Diego

2011