Book Review

Rationality + Consciousness = Free Will, by David Hodgson. New York: Oxford University Press, 2012. Pp. x + 267. H/b £41.99.

David Hodgson's book is a highly ambitious contribution to philosophical debates about free will. These debates are often dominated by the question whether free will is compatible with determinism. Yet it is also a puzzling question how to reconcile free will with *indeterminism*. For libertarians, who think that free will is incompatible with determinism and that we are free, the task is to provide an account that shows how indeterminism is both consistent with and usefully implicated in free decisions. Hodgson tackles the latter project.

Hodgson begins his book in an unorthodox manner by listing his most certainly held beliefs, which include the following: 'Conscious experiences occur' (p. 12) and 'Some conscious experiences change' (p. 14). In the first chapter, Hodgson argues that many of these beliefs are justified and probably true. In chapters two and three, he argues that a presupposition of the claim that such beliefs are justified and probably true is that we are rational. For Hodgson, rationality consists in the ordinary human capacity to make sound judgments about what to believe or do. All rationality (including scientific reasoning) ultimately depends, Hodgson argues, on 'plausible reasoning', which is 'reasoning in which the premises or data do not entail the conclusions by virtue of applicable rules, but rather support them as a matter of reasonable albeit fallible judgment' (p. 7). Hodgson calls this capacity 'ordinary human rationality' (p. 31), and 'instinctive informal rationality' (p. 33). Since plausible reasoning is not fully explainable in terms of rules or laws of nature, Hodgson claims, rationality is not fully explainable in such terms either. In chapters four to six, Hodgson argues that consciousness facilitates rationality — by contributing to plausible reasoning — in a way that does not depend on either rules or laws. So consciousness does the work of explaining rationality. As a result, free decisions can be rational and 'not determined by pre-decision circumstances and laws of nature' (p. 112). It is by thus avoiding the lawlikeness of rational decisions that Hodgson aims to save freedom.

In the remainder of the book, Hodgson places his libertarian theory of free will within a wider context. In chapter seven, he considers whether his account is compatible with current physics, while in chapter eight he develops a picture of how brains might work that is, he maintains, consistent with his

view. In chapter nine, Hodgson considers debates about compatibilism, and he compares his view with other versions of libertarianism. Chapters ten and eleven develop a theory of retribution as well as a theory of moral (and legal) responsibility. In the final chapter, Hodgson considers what everything he has thus far said suggests about the scientific picture of ourselves and the universe, and about competing religious pictures.

A central virtue of Hodgson's book is its focus on the relevance of consciousness for free will. Recent studies indicate that ordinary thinking about free will assigns a central place to consciousness. People tend to consider consciousness as more important to free will than an agent's motivations, values, or character traits, and people attribute free will only to conscious agents (Joshua Shepherd, 'Consciousness, Free Will, and Moral Responsibility: Taking the Folk Seriously', *Philosophical Psychology*, (forthcoming)). This suggests that philosophers should take seriously the question of what positive connection (if any) consciousness has to free will. Hodgson undertakes this task. Whereas others have focused on the apparent disconnect between consciousness and the actual (unconscious) causes of our decisions (e.g. Gregg D. Caruso, *Free Will and Consciousness: A Determinist Account of Illusion and Free Will*, Lanham, MD: Lexington Books, 2012), Hodgson instead develops a view about how consciousness might *contribute* to free will. (For another example of the latter strategy, see Neil Levy, *Consciousness and Moral Responsibility*, Oxford: Oxford University Press, 2014.)

Even so, there is much to disagree with in Hodgson's book. For want of space, I will focus on just two of Hodgson's claims: the claim that (1) plausible reasoning is not fully explainable in terms of rules or laws of nature, which sets the stage for his further claim that (2) consciousness does the explanatory work.

Hodgson rejects the idea that rationality is to be explained by the fact that the processes underpinning human decision-making have proven useful for human survival and reproduction. On that picture, he claims,

> [R]ationality could … be formalised by sufficient specification of the relevant properties of one or more highly rational human brains and of the laws of nature that determine their processes, and/or … by sufficient specification of the computational systems instantiated by these brains and the computational rules of these systems. (p. 49)

Hodgson denies that when there are inconclusive reasons at the top (conscious) level of human cognitive processing, a decision is achieved by rule-determined processes at lower levels, whether these are computational rules or instead laws of nature operating at the level of physical implementation in the brain. Instead, 'conscious processes at the top level make a positive contribution to the resolution of inconclusive reasons, a contribution that is not wholly governed by rule-determined processes at lower levels' (p. 52).

Hodgson's proposal in this regard is the central idea of the book, and he develops it at length in chapters four to six. The suggestion, in brief, is that consciousness plays a role in decision-making by enabling the agent to respond to 'circumstances grasped as wholes, not just to those constituent features that engage with applicable laws or rules' (p. 97).

Hodgson's thesis is intriguing, and he is not alone in suggesting that consciousness contributes in a special way to decision-making. For instance, Hodgson's theory bears a striking resemblance to recent views about prospection, which is the mental simulation of future possibilities for the purpose of guiding action. Although prospection mostly occurs outside of conscious awareness, some have argued that whenever an agent has conflicting thoughts about what to believe or do, her simulated options feed into conscious experience, thereby enabling her to draw on other mental and social resources. This process facilitates resolution of the conflict in a manner that cannot be achieved, the defenders of prospection maintain, solely by implicit algorithmic rules operating at lower levels of cognitive processing, which is precisely Hodgson's claim (cf. Martin E. P. Seligman, Peter Railton, Roy F. Baumeister, and Chandra Sripada 'Navigating into the Future or Driven by the Past: Prospection as an Organizing Principle of Mind', *Perspectives on Psychological Science*, 8 (2013), pp. 119–41).

Nonetheless, I want to suggest that even if consciousness contributes to plausible reasoning in the way that Hodgson or the defenders of prospection claim, it does so in a law-governed way. Let me flesh out this suggestion in terms of causal modelling.

What is missing from Hodgson's picture, I maintain, is causal information, the rules governing how we obtain it, and the stable invariances that it yields (the latter two items are also missing from the view about prospection just outlined). Hodgson might be right that computational (including Bayesian) systems are inadequate to the task of explaining rationality since, for one thing, as Hodgson observes, Bayesianism 'depends on the reasonableness of estimates of prior probabilities, and thus the theorem cannot eliminate the need for informal rationality' (p. 45). Even so, there *is* a formal framework that applies to 'ordinary human rationality', namely the causal modelling framework developed by Judea Pearl (and others). According to Pearl:

> [T]he knowledge we carry in our skulls, be its origin experience, schooling or hearsay, is an invaluable resource in all human activity, and … combining this knowledge with empirical data is the key to scientific enquiry and intelligent behavior. However, in order to be combined with data, our knowledge must first be cast in some formal language, and … the language of probability is not suitable for the task; the bulk of human knowledge is organized around causal, not probabilistic relationships, and the grammar of probability calculus is insufficient for capturing those relationships. Specifically, the building blocks of our scientific and everyday knowledge are elementary facts such as 'mud does not cause rain' and 'symptoms do not cause disease' and those facts, strangely enough, cannot

> be expressed in the vocabulary of probability calculus. (Judea Pearl, 'Bayesianism and Causality, Or, Why I Am Only a Half-Bayesian', in *Foundations of Bayesianism*, ed. David Cornfield and Jon Williamson, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2001, pp. 19–36, p. 27)

To correct for Bayesianism's inadequacies, Pearl develops an alternative framework, in which causal models encode invariant dependency relationships between variables, where the variables represent causal relata. The central idea is that when we consider whether an event is a cause, we ask, 'What if things had been different?' and by answering this question we identify factors whose manipulation would produce different outcomes. This procedure is rule-governed, and it identifies stable invariant relationships among the variables. In turn (because the variables represent causal relata), this information enables us to successfully predict and control aspects of our world, and it regulates our conscious belief-formation and action-planning.

Natural laws, on this view, consist in the invariant relationships that obtain between the variables in a model. The strongest invariances express causal information that is yielded by models of 'local' subsystems of the universe — subsystems that, typically, we are able to isolate and causally interact with. The invariant generalisations governing these subsystems are not derived from 'global' laws — say, from generalisations applying to the universe as a whole. Instead, global laws are constructed from configurations of local subsystems, in a way that first imposes constraints on the variations of the subsystems. The result is that global laws omit causal information, making the content of local laws richer by comparison (cf. Jenann Ismael, 'Against Globalism About Laws', in *The Experimental Side of Modelling*, ed. Bas van Fraassen and Isabelle Peschard, forthcoming).

On this naturalistic picture of ourselves and the world, it might well be the case — as Hodgson maintains — that processes at the conscious level of cognitive processing make an indispensable contribution to decisions, whenever there are competing reasons for believing or doing something. However, even if the contribution is not achieved by rules applying at lower levels of processing, that does not mean that there are no applicable rules or laws at all. The relevant rules are those of causal modelling, and the relevant laws the stable invariances represented in suitable models. Even if consciousness contributes to decision-making in a way that is not governed by global laws, that still leaves room for richer *local* invariances to apply, and such 'little laws' are established by a rule-governed procedure.

There is, consequently, ample reason to doubt Hodgson's central claims that (1) plausible reasoning is not fully explainable in terms of rules or laws, and that (2) consciousness provides the non-law-governed element that does the required explanatory work. Indeed, if the local invariances applying to causal reasoning are consistent with globally deterministic laws (and it is plausible that they are), then Hodgson's libertarian project is undermined.

Whether one finds Hodgson's view, or instead the story I have just outlined, more credible will depend on deep methodological commitments. While it is beyond the scope of a review to settle such meta-philosophical issues, I will just say this. There is at least one fundamental matter about which Hodgson is absolutely right. How we proceed in theorising is ultimately driven by whatever we find it most plausible to believe about various issues, given the aims and goals of our theorising. Hodgson is to be commended for making explicit the commitments driving his own theorising about free will. That is a policy we should all adopt more often.

Department of Philosophy
University of Arizona
213 Social Sciences
1145 E. South Campus Drive
Tucson, AZ 85721-0027
USA
oisin@oisindeery.com
doi:10.1093/mind/fzu168

OISÍN DEERY

