

BOOK REVIEWS

Reasons without Persons: Rationality, Identity, and Time. BRIAN HEDDEN.
New York: Oxford University Press, 2015. viii + 216 p. Cloth \$74.00.

As suggested by the title, this book is in many ways a working out of issues raised by Derek Parfit's *Reasons and Persons*. Parfit's book raised problems for many traditional ethical theories by questioning assumptions about the nature and importance of personal identity. Hedden uses these issues for personal identity to motivate a view he calls "time-slice rationality." Most of the arguments present in this book have appeared in some form or another in his other papers, but this book shows how they come together to form a unified picture of rationality.

Analytic philosophers have traditionally considered "metaphysics and epistemology" to contrast with "value theory." This book is part of a growing body of work that puts central issues in epistemology more clearly on the side of value theory, with metaethics, action theory, and decision theory. The time-slice view that Hedden develops in detail has also been discussed programmatically by Sarah Moss in a few papers that he cites. The Ph.D. program they graduated from at MIT has been quite notable for bridging this supposed divide between epistemology and ethics, and for producing a lot of collaborative work, as well as graduating balanced numbers of men and women. Unfortunately, in this book, only 16 of the 202 works cited in the bibliography have a woman as an author. (This number includes both several women who were in the graduate program with Hedden, and several from other times and places.)

The central idea of this book is to pursue two methodological principles to their logical conclusion. The resulting work is exemplary of a sort of philosophy I quite enjoy. Because of the attempt to take these ideas to their extreme, the resulting conclusions are probably incorrect. However, by following where the ideas lead, Hedden has revealed many interesting ideas and fruitful connections that will likely be of great value, regardless. While a certain sort of theoretical compromise may get one closer to the truth, abstract intellectual exploration can often benefit more from a shocking and stimulating falsehood than from a mundane truth.

The overall picture of rationality put forward in this book is in some ways very minimalistic. The only things that can be rational or

irrational are beliefs, desires, and decisions. We might have thought that there could be rational requirements on updating, learning, reasoning, gathering evidence, and performing actions. However, because these activities take time, requirements on them would run afoul of his first methodological principle, Synchronicity: "How you rationally ought to be at a time directly depends only on your mental states at that time, not on how you (or time-slices psychologically continuous with you) were in the past or will be in the future" (7). The second principle, Impartiality, also blocks synchronic replacements for some of these requirements. Even present evidence of your past or future state is largely irrelevant for current rationality. It does not matter whether you think *you* had/will have a particular set of attitudes but at most whether *someone* did/will.

In other ways, however, the picture of rationality put forward is very maximalistic. Hedden defends a thesis about beliefs and credences known as Uniqueness: "Given a body of total evidence, there is a unique doxastic state that it is rational to be in" (130). He is not quite so committal with regard to desires and preferences, but he expresses some sympathy both for a weaker view on which one's body of total evidence determines what "ultimate preferences" one must have, and for a stronger view on which "there is a unique set of ultimate preferences that it is rational for one to have" independent of one's evidence (159). Supposedly, diachronic principles of rationality can be replaced by these synchronic principles. Facts about the momentary mental state of an agent determine her evidence, which then rationally determines her credences and utilities, which then determine which decisions are rational for her to make. The apparent fact that people ought to learn, update, reason, and act in particular ways over time is then explained by the fact that the evidence someone has normally shifts in particular ways. The standard diachronic rules will usually turn a state that is rational according to the earlier body of evidence to one that is rational according to the later body of evidence.

Chapter 1 introduces Synchronicity and Impartiality, while chapter 2 motivates them through consideration of Parfit cases that problematize the notion of personal identity across time, together with a methodological mentalist internalism about epistemology. Hedden says that the concept of rationality plays three roles with respect to beliefs, preferences, and actions: it guides them, it can be used to evaluate them, and it can often help to predict them from the outside. He says rationality can only play the guiding and predicting role if it supervenes on the mind of the agent. However, he notes that the kind of internalism he needs is compatible with quite externalist pictures of the mind, including the paradigmatically externalist view of Timothy

Williamson that one's evidence is a mental state consisting of all the facts one knows.

Chapters 3 and 4 show how the principles of Synchronicity and Impartiality conflict with traditional rules of rationality like conditionalization, reflection, and analogous rules for desire or preference. He also rehearses some standard problem cases for these rules. The main interest in these chapters is the way that he summarizes the existing literature and discusses principles for desire and preference that are often ignored in the literature on diachronic rules for belief and credence. But these two chapters set up chapters 5, 6, and 7, which attack traditional arguments in favor of these rules, and then chapters 8 and 9, which present his replacement for these principles in the form of rational uniqueness principles, mentioned above. Chapter 10 contains a discussion of apparent norms on reasoning, gathering evidence, and acting over time, and chapter 11 is a summary and conclusion.

The argument of chapters 5, 6, and 7 is central to the book. Chapter 5 notes that most arguments for diachronic rules take the form of noting that a violation of the rule entails "diachronic tragedy." That is, at t_1 the person has options A_1 and B_1 and prefers to do A_1 (regardless of what she will do at t_2), and at t_2 the person has options A_2 and B_2 and prefers to do A_2 (regardless of what she did at t_1), but at every time, the person prefers the sequence $\langle B_1, B_2 \rangle$ to the sequence $\langle A_1, A_2 \rangle$. To avoid the tragedy, one must follow the diachronic rule. Hedden notes that these arguments are given not just for diachronic rules, but also for some principles such as the ones stating that credences must always be numerically precise, that preferences must always be transitive, and that there are no incomparable values. These other principles seem out of place at this point, but he recommends denying them for more plausible versions of his uniqueness theses in chapters 8 and 9, so he wants the argument of chapters 6 and 7 to permit these violations as well.

Chapter 6 starts with the seemingly unrelated observation that to evaluate the rationality of an option selected, we must know what options are available to the agent—it is no argument against X to note that Y is better unless Y was an option. Hedden argues that since the set of options determines rationality, it must supervene on the mental state of the agent. Thus, the options must not just be the acts the agent *can* in fact do (since two agents in identical mental states might come across similar creeks while hiking, and one of them can ford her creek, while the other cannot, because of some non-mental difference in their situation). Instead, he argues that the options must be the *decisions* that the agent can make, which we might identify with "tryings."

(There is an interesting discussion on p. 104 of Timothy Williamson's argument that knowledge is important to explaining the success of actions over time. Hedden thinks that only momentary decisions can be rational, though extended actions can be explained.) Chapter 7 brings this back to the arguments from chapter 5—since there is no point at which the agent can decide on the *sequence* $\langle A_1, A_2 \rangle$ rather than the sequence $\langle B_1, B_2 \rangle$, the sequences cannot be said to be irrational. He compares these diachronic tragedies to the interpersonal tragedies of the Prisoner's Dilemma. If at t_1 the agent can form a causally efficacious *intention* to do B_2 at t_2 then she should, but the lack of this ability is no more irrational than is one's inability to cause one's partner to cooperate in the Prisoner's Dilemma.

Chapter 8 is the longest in the book (nearly twice as long as any other), and I cannot do justice to its discussion of uniqueness principles in this short review. It considers cases from Parfit, the notion of evidence, imprecise probabilities, impractical preferences, and David Lewis's "desire as belief" thesis. I have some worries about whether the approach to the problem of forgotten evidence might undermine other parts. But the central observation that uniqueness principles can replace diachronic and interpersonal ones is important. Hedden compares this to a claim by Niko Kolodny that supposed norms of consistency can be replaced by norms to follow the evidence, if the evidence never points in inconsistent ways.

Chapter 9 then aims to replace principles of deference to one's own future beliefs and desires. As Hedden notes, these "reflection principles" can be phrased in purely synchronic terms, saying that one should not disagree with what one thinks one's future, better-informed self will believe or desire. He subsumes these principles under general principles of deference to experts—if one believes that *any* better-informed, perfectly rational person will believe p , or will prefer X to Y , then it should not matter whether that person is one's future self or not. The discussion of "preference deference" principles is of note not just because of the catchy term, but because these principles are radically under-discussed in the literature. I would have liked to see even more of this chapter, engaging with much of the literature on disagreement and testimony (which would have provided opportunities to engage with many other philosophers who are not addressed).

A point I find extremely important for understanding the implications of this overall view and whether we should accept it comes up in the last six paragraphs in the book. Hedden notes that his view assumes that time-slices of persons are atomic, while successive time-slices are to be treated just as any other group of agents. But he

mentions examples from other philosophers suggesting that even the time-slices are not as atomic as we might think. Peter Godfrey-Smith points out that in an octopus, the eight legs each have their own nerve bundle that is almost as strong as the central brain, and they exhibit relatively coordinated, but disunified, behavior. David Lewis and Jerry Fodor have given reasons for thinking that human minds also exhibit some sort of fragmentation or modularity. And split-brain patients (discussed by Parfit, among others) are perhaps the most famous case that blur the line between one and two selves.

Consider a split-brain patient presented with a version of the Prisoner's Dilemma. Each hand is presented with a button. If the button is pressed, that hand will receive a mildly pleasant massage, while a mosquito will be released to bite the other hand. From its own hedonistic perspective, each half of the brain prefers a slight massage regardless of whether it gets a mosquito bite, and prefers no mosquito bite regardless of whether it gets the massage. But overall, it prefers neither to both. On a strict construal of Hedden's considerations, it seems there is an interpersonal tragedy here, where each half brain would be rational to push its button, and would unfortunately end up with a mosquito bite and a massage.

There are many natural responses here. One might say that the two half brains in fact care about each other's pain as their own. Or one might say that the two half brains share a body and, thus, have some rational obligation to cooperate. Or one might hope that even with a severed corpus callosum, there are other connections (perhaps through the spinal cord or physical closeness of the hands) that allow each half brain to partially control the other hand as well as its own. However, all of these considerations exist to some extent between one time-slice and another, and to more attenuated degrees between a person and her spouse, her best friend, her neighbor, her compatriot, or even a perfect stranger who happens to inhabit the same planet.

On p. 21, Hedden considers David Gauthier's accusation that utilitarians treat all of humanity as a kind of "super-person," as well as Parfit's response that he is *disintegrating* persons rather than conflating them. But I think it is instructive to consider the similarities between these two apparently diametrically opposed ideas. In *The Selfish Gene*, Richard Dawkins popularized the idea (discussed earlier by George Williams and Richard Lewontin) that the unit of selection in biology is not the organism but rather the gene. This dissolution of the individual leads to evolutionary behavior that looks a lot like selection at the level of the group. A gene for interpersonal cooperation may hurt the individual who has it but help enough other individuals with copies of that gene to make up for it. And while biology is not rationality,

some of the rules for rational cooperation that Gauthier proposes resemble the ethical views defended by Parfit in ways that gene selection can resemble group selection.

Whatever is present at the level of the individual time-slice to rationalize coherence may be present to an even greater degree at some sub-individual level, and to a lesser degree across time-slices and across persons. The counterexamples to diachronic principles may be cases in which this connection is suppressed due to forgetting or change in character. Cases of expert testimony or peer disagreement that seem to call for some sort of interpersonal coherence may be cases in which the connection exists between apparently separate individuals.

Hedden's book works out motivations and arguments for a view that denies these connections and is based on an atomized and individualistic picture. I find this view fascinating, primarily as an enterprise to imitate in mirror image. Rather than making us feel farther from our past and future selves and treat them more impersonally, I think Parfit's cases should make us feel closer to distant people and treat them more personally. Where Hedden goes for radical uniqueness, I go for radical permissivism. Where he derives expert and reflection principles from a universal shared prior, I would derive them from a general commitment to cohere with agents that (to some degree) share your values, or with whom you can coordinate. Whether you share my views or Hedden's or fall somewhere in between, there is a lot to learn from considering this project.

KENNY EASWARAN

Texas A&M University