

Mendola, Joseph. *Human Interests: or Ethics for Physicalists*.

Oxford: Oxford University Press, 2014. Pp. 432. \$99.00 (cloth).

In *Human Interests: or Ethics for Physicalists*, Joseph Mendola develops and defends a new form of direct consequentialism, rooted in a simple desire-based account of wellbeing, which is compatible with Physicalism. He calls this theory ‘Material Morality’ (or ‘MM’). Initially, it might be hard to imagine how a plausible moral theory could result from joining these three views, as each is susceptible to familiar objections: Direct consequentialism holds that it is sometimes required that we perform intuitively immoral actions like murder. Simple desire-based accounts of wellbeing state that all and only the things that we actually desire are good for us, when intuitively, we can actually desire things that are bad for us (and we can fail to desire things that are good for us). Physicalism states that all that there is in the world is particles and void, and it is not clear how we can generate the normativity that is characteristic of moral properties from mere particles and void. What makes this book so exciting is that Mendola offers new answers to these familiar objections, and consequently succeeds at offering a plausible moral theory composed of these apparently implausible components.

The primary goal of the book, however, is not just to convince us that MM is plausible. Mendola’s more ambitious thesis is that given that Physicalism is true, MM is the only possible moral theory. His argument spans most of the book, and covers an impressive number of topics in metaethics and normative ethics (as well as a few topics in action

theory and metaphysics). Anyone working in these areas is sure to find this book to be worthy of their time and attention.

Mendola begins with an exploration of the various ways in which the physical nature of reality constrains which types of moral theories could actually be true. First, and most obviously, any moral theory that posits non-natural properties of goodness or rightness simply cannot be true in our physical world (2). Surprisingly, Mendola concedes to the familiar objection that natural properties cannot deliver full normativity (characterized by the ability to immediately terminate justificatory regresses and to generate external reasons for acting). Since non-natural properties do not exist in our world, then neither does full normativity. Instead, says Mendola, we must settle for ‘quasi-normative’ properties which are similar enough to their non-natural counterparts to offer an end to justificatory regresses and to generate internal reasons (95-100).

The bulk of Chapters Two and Three, though, focus on a more interesting implication of Physicalism: in many cases, it renders indeterminate what options are open to an agent at a time. This indeterminacy is metaphysical, not epistemic. According to Physicalism, the world is either fully deterministic or very nearly so, and thus we cannot do otherwise than we in fact do. To preserve the commonsense stance that we do have options (for instance, that I could have ordered a beer instead of the wine with dinner), Mendola favors the Conditional Analysis of Ability, which says that an agent “P has an alternative y at t if and only if there is something such that if P were to try to do it at t, then P would take alternative y at t” (40). If I had tried to order a beer, I would have taken that option.

When assessing the truth of counterfactuals like the one in the Conditional Analysis of Ability, we need to consider what occurs in ‘nearby’ possible worlds that are different enough from the actual world for the antecedent of the conditional to be true, but still very similar to the actual world. Preserving similarity between the actual world and ‘nearby’ ones requires selecting a set of ‘background conditions’ that are held fixed across those worlds. However, in many cases we will have a choice about which set of background conditions to hold fixed, and holding different sets fixed can result in different truth values for the conditional. In some cases, there is no principled way to select among the competing sets of background conditions, and there is consequently no fixed truth value for the related counterfactual. Thus, in some cases there really is no fact of the matter on whether a certain option is open to an agent at a time.

Indeterminacy also plays an important role in several of Mendola’s arguments in Part II, where he does most of his work in metaethics and normative ethics. For instance, Mendola appeals to the indeterminacy of counterfactuals in his argument against informed desire accounts of wellbeing (160-5). According to those theories, what is good for an agent is what they would want if they were fully informed. Clearly we have a great deal of choice in deciding which features of the actual world we will hold fixed when we assess the truth of this counterfactual, and clearly which features we hold fixed will affect the truth value of the counterfactual. In many cases, there is consequently no fact of the matter of what we would want if we were fully informed, so informed desire accounts are unworkable in reality.

In Chapter Five, Mendola explains why he prefers a simple desire based account of wellbeing. He argues that hedonistic and objectivist accounts of wellbeing either require objectionably ‘fancy’ metaphysics, or collapse into a form of desire theory. Thus, the only plausible account of wellbeing that could be true in our physical world is a simple desire based account. I will discuss these arguments in more detail below.

There are some unique features of Mendola’s theory of wellbeing that help it to avoid the familiar objections to simple desire theories. He prefers a broadly motivational account of desire wherein to desire that S is to be disposed to bring about S (172). According to his view, what an agent most desires will be whatever she would most prefer if you were to present her a lengthy (but finite) series of pairs of options. In this way, even if Arthur currently believes that what would satisfy his strongest desire is to stay seated at a table (where, unbeknownst to him, a piano is about to fall), he would probably rank leaving the table as preferable to being crushed by a piano if presented with those hypothetical options. So, what he really wants is to leave the table (205). Mendola’s account of wellbeing can thus explain why staying at the table would not really promote Arthur’s wellbeing, even though Arthur might insist that staying is what he most wants to do at the moment.

Having settled on an account of personal good, Mendola spends Chapters Eight and Nine specifying the mechanisms of MM that prescribe how we ought to distribute wellbeing in interpersonal cases. One might expect Mendola to defend a kind of utilitarianism that

seeks to maximize overall wellbeing for everyone affected by the action. However, he rejects such a view on the grounds that we can rarely assign cardinal values of wellbeing to actions, and as such, it is impossible to sum them together in the way required for maximizing strategies (226-8). What we can do, however, is offer an ordinal ranking of preferences that are shared by nearly everyone. Mendola thus proposes that we distribute wellbeing by way of the Leximin Desire Principle. Roughly, what we should do is maximize the wellbeing of the worst-off.

The second component of MM is what Mendola calls Multiple Act Consequentialism, or MAC. At every moment, an agent has the option of participating in a variety of group actions. Mendola's theory could not give us guidance unless it provides a set of principles that prescribe when we should join group acts we are not currently a part of, and when we should (and should not) defect from group acts that we are currently pursuing. Roughly, Mendola's answer is that we must defect from group acts that are harmful, we must join the most beneficent group acts, and we are forbidden from defecting from a beneficent group act unless our defection brings about more utility than the entire group act (257).

MAC is Mendola's response to the familiar objection to direct forms of consequentialism. According to the principles specified by MAC, one should not murder unless doing so has a greater benefit than the overall benefits of the extended group act of not murdering. That is a high bar to meet, so MM delivers a near universal prohibition against murder. Similar reasons speak against lying, stealing, etc. So while MM is a form

of direct consequentialism, it retains the intuitive deontic constraints of commonsense morality (336-8).

I have merely gestured at some of the highlights of the book. Unfortunately, with a book that is as rich in arguments as this one, any summary is necessarily going to have to skip over something of interest. I would be remiss if I did not mention Chapter 10, where Mendola argues that the only viable forms of “the utilitarian, contractarian, and golden rule traditions in both its Kantian and non-Kantian varieties, all find their appropriate real world development in [MM]” (331). This chapter is worthy of careful study. Among the other arguments I could not summarize are: an argument that momentary ‘atomic’ agents are the true loci of wellbeing (Chapter 8), that there is no self over and above a temporally extended group action (Chapter 8), that MM can explain the value of many of the virtues (Chapter 12), and that MM is not ‘too demanding’ (Chapter 12).

As I hope is clear from my discussion, the overarching argument for MM is a thing of beauty and elegance; it is like a timepiece constructed out of many intricate, interlocking pieces. The result, however, is a very delicate moral theory. Should we find fault with any of the many claims or arguments that constitute the argument for MM, then the theory might grind to a halt. In what follows, I will suggest a few worries about the theory of value that Mendola argues for in Chapters 5 – 7.

Traditionally, theories of wellbeing fall into three broad types: hedonistic, desire-based, and objectivist. Theories of each type will largely agree with one another about what

particular things are good (they will all extol pleasure, desire satisfaction, friendship, health, the cultivation of talents, etc). Given that they mostly agree on what things are good, the main difference between the three types of theories is their explanation for what makes those things good (their ‘rationale,’ as Mendola calls it).

Mendola rejects hedonistic and objectivist views on account of their rationales. He claims that they either require metaphysical buttressing that is not possible in our physical world, or that they are really, deep down, just desire-based theories of wellbeing. I wish to suggest quite the opposite. Either Mendola’s simple desire-based view rests on an implausible rationale that is not sufficient for generating even quasi-normativity, or it is really, deep down a hedonistic or objectivist theory of wellbeing. I will illustrate my concern with an example.

Suppose that in the not too distant future it becomes possible to purchase a robot butler whose primary tasks involve the cleaning of the house. The programming for the robot stipulates an ordinal ranking of its tasks—it should dust before it vacuums, it should vacuum before it mops, etc. On Mendola’s view, we are committed to saying that the robot has preferences; since preferences are nothing more than ordinal rankings of what one would do in different hypothetical situations. If the robot had just finished dusting, then it would prefer vacuuming to mopping. And since on Mendola’s view desiring that S is just preferring that S, we would also be committed to saying that the robot has desires. This is counterintuitive. While we often talk as if non-living apparatuses have desires (we might say, for instance, that a struggling car engine ‘wants to start’), this kind of talk is

essentially metaphorical. Cars, and robot butlers (at least of the simple sort described above) do not have mental lives; they do not have desires. On Mendola's account of desire, however, it is *literally* true that our robot butler wants to dust.

There are stranger implications of his view. Because he offers a desire-based theory of wellbeing, Mendola would also be committed to saying that satisfying the robot's desires promotes its wellbeing, while frustrating its desires detracts from its wellbeing. Perhaps Mendola would be happy to concede all this. After all, we cannot meaningfully compare the robot's ordinal preference rankings with our own, so we cannot share preferences with a robot in the way that is necessary for us to judge how to weigh its wellbeing in our use of the Leximin Desire Principle (for similar reasons, Mendola's view highly discounts the wellbeing of non-human animals). So while robotic butlers have morally relevant desires, we could frustrate these desires even for the most trivial human interests.

That reply handles most cases, but what should we say in a case where there are no contrary human interests? Suppose I am considering refurbishing the floors in my house. I can either install new carpet, or I can refinish the hardwood floors that are beneath the old carpet. As luck should have it both procedures are equally time-consuming and expensive, and I would find either result equally aesthetically pleasing. I do not have a preference either way, so my wellbeing would be equally promoted whatever I do. I am about to flip a coin to decide how to proceed when I remember that I own a robot butler who prefers vacuuming to mopping. Refinishing the wood would make the robot worst

off than replacing the carpet, so the Leximin Desire Principle requires that I replace the carpet for the sake of my robot butler.

When we focus on what Mendola means by ‘desire,’ and recognize that even a mindless robot can have desires of this kind, it undermines the plausibility of the rationale for Mendola’s desire-based theory. Why is it good for us to have our preferences satisfied when we understand desires in this purely motivational way that Mendola favors? A more plausible rationale for the goodness of desire satisfaction would either appeal to the pleasure we get from having our desires satisfied (and thus be a hedonistic rationale), or to the goodness of the objects of our desires (and thus be an objectivist rationale).

Mendola’s view states that pleasure is good because we share a strong preference for having pleasure. When I in a pleasurable state, then, what makes that state good for me is not the pleasantness of the state—nor is it the feeling I get when I have my desire for pleasure satisfied. For Mendola, pleasure is thus no more a goodness-making property than is finding a parking spot close to the movie theater. Of course, pleasure is *better* than finding a close parking spot (because we rank our preference for pleasure quite highly); it is not however, special. On Mendola’s view, it is just one among so many other preferences, which intuitively, is not the case.

Brian Coffey

Grand Valley State University