Kenny Jones

11/17/14

On *Elbow Room*

In his book *Elbow Room*, Daniel Dennett makes a compatibilist argument that determinism does not render our current moral system obsolete. He believes that in these circumstances people can still be held morally responsible for their actions and that they can have free will in the sense that they have the power to decide their courses of actions, are in control of themselves, and are able to take responsibility for their projects and deeds (169). While Dennett makes compelling arguments for how it is possible to keep our current moral system with determinism, I believe that his account falls short in two places. It seems to me that his explanation of morally responsible agents necessitates that they are not in full control of their actions when they do wrong, which contradicts his definition of a free agent, and also, that he contorts and manipulates the concepts of free will and moral responsibility in such a way that they no longer resemble their original definitions that were based on metaphysical concepts.

Dennett believes that "the sort of free will that any responsible, dignified, moral agent must have" does not have to come from contra-causal freedom or agent causation (153). Instead, he believes that a moral system that holds agents who have free will responsible could exist in a determined world. He thinks that the line of moral thought supported by Kant, following from Socrates, that only the good actions are responsible ones, is misguided as it treats fallible humans as perfect beings (156-157). Therefore, he says that his purpose is to find "elbow room for us sinners in between the saints and the monsters"; that is, to find out when and how an agent can be held responsible for a wrong action (157).

Dennett tackles this problem by approaching the subject with an eye on efficacy. In an ideal world, everyone would perform perfect moral action, and there would be no need for punishment, but as "we are no angels" (158), there remains the problem of dealing with these wrong actions. He argues that deterrence, in the form of laws and punishment, provides an effective way of dealing with these incidents of criminal activity. Dennett justifies this position by asserting that all crimes are not done out of pure ignorance, but that some are done out of calculated deliberation (159). Most of the time, we rational creatures can see that punishment is to be avoided, but sometimes criminals believe that "crime does pay" (159), and this is what triggers wrong moral action. Dennett believes that an "optimal institution" of enforcement "will be one in which a certain amount of law-breaking, apprehension, and conviction is 'tolerated'" (160), for past a certain point there are diminishing returns on increasing punishment.

Dennett also posits that a key factor of an enforcement system would have to include publicity (160), as it is only when an agent knows the law that he can adjust his behavior accordingly. He warns that this train of thought can lead to the Kantian argument made earlier: that agents should not be blamed for their wrong actions as these are the just the result of simple oversight or ignorance on the part of the agent, and do not stem from something that would merit real blame. However, Dennett rejects this notion as it would completely wipe out any effect of deterrence. Rather, he wants to create a "constitute – a class of legally culpable agents whose subsequent liability to punishment maintains the credibility of the sanctions of the laws" (162). He argues that while on an individual case by case basis there might be excusable reasons to exempt a particular agent's actions, a general set of rules should not be overly bent to accommodate these cases. That "holding people responsible is the best game in town" because it is the most socially efficacious solution. Moreover, Dennett believes this system can be fairly

justified: “as long as the risk was taken in full knowledge of the consequences of the loss, the agent can hardly complain that the sanctions now imposed are unfairly applied to him” (162).

Although Dennett makes a strong argument that we would not have to abandon moral responsibility in a determined world, I believe that parts of his argument seem to leave the agent in a state in which he has no power or control over his actions. Dennett makes the point that a morally bad action is not always the case of faulty causal reasoning leading the agent astray. Rather, he explains that for the “best possible designs” of memory organization, there must be some level of “arbitrariness and wise risk taking” (164). I cannot disagree with this point, but I will assert that this notion goes against our common thoughts on free will, and Dennett’s own definition of it. For an agent to have free will, Dennett requires that the agent must “decide [their] courses of action…wisely” (169), but doesn’t the arbitrary nature of these “gambles” seem to undercut the agent’s ability for self-control?

Suppose that there is a perfectly rational agent who always wants to perform the most morally good action. For the vast majority of his life, he only performs good actions by managing to correctly and efficiently choose the best decision using these shortcuts that Dennett highlights. However, at times he will be faced in an unclear moral situation in which he will have to choose between two very difficult options. Dennett would argue that even if this agent had all of the necessary information to make the correct decision in his mind, he could still make a wrong guess occasionally, and be left to face the consequences of his decision. While I do not disagree that this agent should be held responsible for his choice in this scenario, I am hesitant to say that he was a genuine free agent who had meaningful self-control. From the example, we know that the agent had both a wish to do the right thing and all of the information necessary to make the correct choice; nevertheless, the agent still has the possibility to commit the morally

wrong action. This does not seem like true self-control, but rather, that the agent's actions are dictated by unconscious forces that he cannot exercise power over. Dennett defends his position by appealing to the notion that all decision making processes have to stop at some point, and these stopping points can be arbitrary. He cites the examples of a military draft that is used to speed up a process that would otherwise be too time and resource intensive (165). However, it is unclear how this example would show that the agent still has meaningful control over all of his actions. If the agent does not have the power to go *all the way through*, if they so choose, with these "intractable decision problems", then the agent is neither in total self-control of his actions nor completely free in our typical meaning of term.

Building off of this example of self-control, it seems to me that Dennett has a way of distorting the notions of topics relating to free will in a disingenuous way. Dennett is very aware and vocal of the fact that he is not interested in the typical metaphysical conceptions of free will that stem from what he thinks are misleading philosophical ruminations. Instead of some abstract notion of free will, Dennett is more interested in examining how we could have most of the benefits of this idealized free will in our practical world. However, I believe that by asserting this, and then attempting to show that we have *free will* in a practical way, he changes the meaning of the term fundamentally.

Dennett knows that his version of free will does not give the agent all that he wants; it is, in some sense, just an imitation of the real thing. As seen earlier in the paper, Dennett treats self-control in just this way. Although perhaps not practically possible, an agent would like the ability to perfectly control his actions as one of the prerequisites for free will, not just the partial control that Dennett argues we have. In a similar vein, Dennett also argues that we do not have to have a perfect Kantian will to be a moral agent and that it is impossible for there to be a completely self-

made self that is responsible for its own character, but that we can have "worthy approximations of these imaginary absolutes"(156-157). But in stating this, Dennett adds an asterisk of sorts to these terms. There are no longer moral agents, just Dennett's moral agents who adhere to these guidelines. Agents no longer have free will, they just have Dennett's version of free will. Dennett creates new concepts that would work inside of a determined world, but in no way can he keep *free will* in the larger sense, because this is exactly what he is trying to bring down.

I will approach this argument in another way by examining Dennett's proposal to continue to hold agents morally responsibility. From Dennett's line of reasoning, it is difficult to provide a counter-argument that it would not be efficacious to hold agents responsible for their actions and to punish them accordingly. However, this theory seems to have as its primary justification: we should keep the system because it is the *best way things are designed to work*. The "worthy approximations" seem to be developed out of their social function, which runs counter-intuitive to our usual understanding. We commonly believe we *can* be held responsible *because* we are free and responsible agents; we do not use the fact that we *are* held responsible to justify our freedom. Dennett has taken to the second route because he attempts to ground his new compelling conceptions of these terms in an explanation of societal usefulness, but in doing so he redefines them in such a way that they lose their meaning.

But one cannot just abandon a concept's metaphysical roots because one believes them impossible. On page 168 Dennett argues that, "having good reasons for wanting free will is not, of course, having good reasons for believing one has free will. It seems to be, however, that having good reasons for wanting free will *is* having good reasons for trying to get oneself to believe one has it". But this process is not as simple as Dennett makes it out to be. If agent A has a certain idea of free will that is based on some 'impossible metaphysically notion', he cannot

simply make himself believe that he is now a free agent because he has something *close to* free will. Agent A can rationalize in his mind that he may not need free will, that, in fact, this other free will may in some ways be more appealing than the typical free will, but he cannot simply discard his previous associations of this ingrained concept at the drop of the button. Dennett is trying to redefine words with meanings that run counter to their original definitions.

Instead of taking a compatibilist stance, I believe that Dennett's argument would resonate more if he cast it in a hard-incompatibilist light. He could structure his book similarly to Pereboom's *Living without Free Will*: argue that not that much would change in a determined world, but that people would have to accept living under these new concepts. I do not think Dennett's account can truly be called compatibilist, as it denies the existence of metaphysical freedom, and only inserts imposter substitutes to fill the gap, but this does not make his arguments fundamentally wrong or uninteresting. If he took a hard-incompatibilist stance I think his account would be truly compelling, and he could make the argument that while you cannot have *true* free will, as the concept is impossible, agents would still be left with something almost as good. If Dennett were to add the necessary qualifiers to the terms he is at once trying to denounce and affirm, then his argument would be greatly strengthened. There is nothing wrong with accepting or promoting free will with an asterisk; Dennett just can't leave out that the asterisk is attached.

Works Cited:

Dennett, D. C. *Elbow Room: The Varieties of Free Will worth Wanting*. Cambridge, MA: MIT, 1984. Print.

Bibliography:

Pereboom, Derk. “Living without Free Will: The case for Hard Incompatibilism.” In *The Oxford Handbook of Free Will,* ed. Robert Kane (OUP, 2002), 489-505