FREE WILL AND PREACTIONS

JAKUB BOZYDAR WISNIEWSKI[*]

METAPHYSICAL LIBERTARIANISM (hereafter ML) is the doctrine that human beings possess free will, that free will is incompatible with determinism, and that determinism is false. Its nomenclatural affinity with political and economic libertarianism (hereafter PEL) is by no means accidental, since, as I am going to argue, the viability of the latter depends on the viability of the former.

I believe that no argument is needed to convince the readers that the so-called "hard determinism," which rules out free will, and hence also independent personal choice, is incompatible with PEL. On the other hand, the "soft" variety of determinism, known under the name "compatibilism," is oftentimes claimed to be reconcilable with PEL. According to compatibilism, the assumption that every event (including every event of personal choice) is causally necessitated by antecedent events, and the resulting conclusion that nobody could ever have chosen otherwise than he in fact did, are perfectly compatible with laissez-faire.

I think this is mistaken—I remain convinced that as soon as one grants that every human decision can be traced back to factors beyond one's control (e.g., genetic makeup, environmental influences, personal upbringing etc.), the notions of sovereign choice and personal liberty become empty. For instance, the above concession enables the so-called "luck egalitarians" to claim that an adult man who remains unemployed on a free market ended up in such a situation involuntarily, since, e.g., his lack of appropriate competences, and his lack of willingness to gain any, are determined by genetic and environmental factors, over which he had no control—

[*]Jakub Bozydar Wisniewski (j.wisniewski.1@research.gla.ac.uk) is a philosophy graduate from Cambridge University, currently doing postgraduate research at the University of Glasgow.

consequently, his predicament should count as an instance of "brute luck" rather than "option luck," and thus should be offset by welfare subsidies.[1]

The internal logic (or lack thereof) and exact political implications of the above scenario need not concern us here, however. In this paper, I shall chiefly focus on a different, even more thoroughgoing and illustrative example.

In a recent issue of *Analysis*, Saul Smilansky argued for the thesis that the acceptance of compatibilism implies the endorsement of the practice of prepunishment, i.e., the practice of punishing people before their commission of a crime (Smilansky 2007). Such a procedure, in my opinion, is as inimical to freedom of action, self-ownership and other crucial tenets of PEL as anything can be. Whether it follows from compatibilism, however, is a contentious issue— in the ensuing debate, Stephen Kearns (2008) and Helen Beebee (2008) suggested that compatibilism can resist this disturbing conclusion. Smilansky then responded to their criticisms (Smilansky 2008a, 2008b).

By further exploring some of the issues touched upon in the aforementioned debate, I shall make a claim that considerations of prepunishment, as well as related advance actions, which I shall collectively call "preactions," not so much reveal and underscore the radical consequences of compatibilism, but rather, firstly, threaten its collapse into hard determinism, and secondly, cast a shadow of suspicion on determinism itself, thus opening some new, promising avenues for ML, and hence also securing the metaphysical viability of PEL.

Let us start from analyzing the claim made by Beebee that, given the possibility of time travel, prepunishment is as compatible with libertarian free will as it is with compatibilism; Beebee argues that, since travelling into the future enables us to establish conclusively that a given person has committed a crime, what really lends plausibility to the notion of prepunishment is not determinism but perfect predictability. In his reply, Smilansky is rightly suspicious of taking time travel in this context seriously, but he does not identify the source of his reservations; I believe that the (correct) reason for having them is that ML is incompatible with the idea of a timelessly and changelessly complete series of events, akin to the McTaggartian B-series, where the time-slice corresponding to any of these events can be accessed from any other time-slice by leaping over the temporal interval that lies between the two.

---

[1] For an elaboration of the distinction between brute luck and option luck, see Dworkin (2000).

According to the libertarian worldview, ontological constituents of the present neither contain nor extend into their definite future histories—the life of a pre-autumn 1888 unidentified Londoner need not extend into the life of post-autumn 1888 Jack the Ripper, nor does the pre-conclave Joseph Ratzinger contain the post-conclave Benedict XVI. In short, at any given moment t-2, there exists no future moment t to be visited, since the latter can emerge only from events taking place at the immediately preceding moment t-1. Hence, ML does not admit of perfect predictability,[2] be it through time travel or Laplacean calculations.

Having said the above contra Beebee, I will now turn to questioning certain claims made by Smilansky; I do not understand, for instance, how he can decouple prepunishment from perfect predictability and say things like: "the reason why we may punish at t0 [t1 being the time when the crime takes place] is not predictability but that the crime, in a sense, is already there" (Smilansky 2008b: 261). Surely, we cannot know whether the crime is already there unless our predictions about the future are perfect. Given determinism, the pre-existence of crime at t0 is not metaphysically problematic, but unless we solve the epistemological problem of conclusively detecting its presence, we are not justified in prepunishing its perpetrator.

Smilansky's suggestion is that "since this person [the perpetrator] is highly reliable, we know beyond a reasonable doubt that he is going to commit the offence" (Smilansky 2008a: 254). But do we really? Is high reliability enough for prepunishment? Let us contrast this case with that of another kind of preaction, namely, prereward. Imagine a running competition between a team of professional sprinters and a team of wheelchair-bound invalids; it appears clear to me that we know beyond a reasonable doubt that the sprinters are going to win, but is that a reason to reward them prior to the competition? And if so, what is the reason of organizing the competition in the first place? I believe that the answer is that organizers of sporting events, as well as court judges, generally recognize and respect the epistemic difference between high reliability and perfect predictability, according to which preactions are mandated only by the latter. If this difference is illusory and compatibilism is true, prerewards are just as justifiable as prepunishments. If, on the other hand, this difference is real and compatibilism is true, compatibilists unequipped with perfect predictive powers turn out to practice prepunishment not because in any given case the crime is in a sense simultaneous with the intention to commit it, but because of the fear that the crime is highly likely to happen in a more or less specified

---

[2]Even with regard to non-human (and thus presumably non-free-willed) elements of reality, since it cannot be known whether they will not come into interaction with human beings at any specific point in the future.

future (in which case what they practice is not really *pre*punishment, but ordinary *pre*emption).

So let us suppose that some ingenious compatibilist judges come up with a prototype of a Laplacean calculator (a device perhaps philosophically easier to conceive of than a Wellsian time machine). They put into it the initial data of the universe and the whole course of cosmic events is revealed before their eyes. The crucial question now is: how should they interpret and utilize what they see? That, of course, depends on what they see, but I am inclined to think that the obtained results admit of more than one interpretation. If they learn that they are determined to prepunish an individual X at $t0$ for committing a crime that X nonetheless succeeds to commit at $t1$, then it seems that in this case their act has a purely retributive character. Associating *pre*punishment with a paradigmatically *backward*-looking function of penal practice appears, to me at least, very bizarre or even confused, so I shall return to the above case a bit later and suggest that with regard to it other solutions should be looked for, solutions based on prophylactics or prevention rather than retribution.

But for the time being, let us turn to another worry, raised by Kearns, namely, the suggestion that X committed a crime precisely because he had been prepunished. It seems that even knowing the entire history of the actual world cannot help the judges determine whether in any given scenario this is the case or not. For instance, in the scenario mentioned above, the calculator could show that the criminal intention stayed with X ever since he had formed it and that prepunishment did not alter its intensity in any way; this, however, does not mean that prepunishment had no influence on it, since it is possible that the intention in question could have faded away at some point after $t0$ had X not been prepunished at $t0$ (of course, given determinism, it is a metaphysical "could" rather than a nomological "could," but the problem is that we do not know whether what precludes it from being nomologically actualizable is X's character or the actions of the judges).

Similarly, if the judges discover that they are determined to prepunish an individual Y at $t0$ for committing a crime at $t1$ which, as it eventually turns out, Y does not even attempt to commit, then it is not clear whether the act of prepunishment deterred Y from criminal activity or constituted a redundant and unjust response to an ultimately non-existent threat (presumably, administering the prepunishment in this scenario is supposed to be justified by reference to the fact that the relevant criminal intention persists until $t0$, but that does not preclude the possibility that it could have faded away sometime in the interval between $t0$ and $t1$ without the help of any punitive incentive).

The solution to the above worries appears to me to be the introduction of a Laplacean *super*calculator, capable not only of modeling the entire history of the actual universe, but also of accomplishing the same feat with regard to various possible universes.[3] Its user, having looked at the Laplacean model of the actual world, could ask the system to subtract from it all the factors that contributed to the occurrence of the act of prepunishment at $t0$ and then run the modeling process anew. The counterfactual simulation thus obtained would demonstrate how the life of the putative criminal would have looked like had he not been prepunished. For instance, with regard to the person labeled Y in one of the aforementioned cases, it could turn out that even though he harbors the relevant criminal intention up until or even after $t0$, this intention evaporates before $t1$, in which case no deterrent is needed. Things might unfold similarly with regard to X, in which case the prepunitive actions taken against him in the actual world (the one in which the judges use an ordinary, not a *super* calculator) appear to serve as an irritant, solidifying X in his criminal inclinations. Given such outcomes, the judges would do best to remain inactive in both cases.

This, I believe, puts Kearns's worry about the possible initiation of factually illegitimate preactions to rest. So let us now return to the case in which, regardless of whether the prepunishment takes place or not, the criminal carries out his unlawful plans. As I already pointed out before, here any potential prepunitive action seems to me to acquire a purely retributive character. So far so good; but then it becomes puzzling how Smilansky can reconcile such retributivism with his insistence that, as soon as it is determined that the relevant criminal intention will not vanish, "there seems to be no point, from a compatibilist perspective, for waiting" (Smilansky 2008a: 255). If determinism is true (minus quantum uncertainties), then presumably it was determined from the beginning of the universe that the relevant criminal intention will be formed and carried out (provided that no preactive interferences take place). Thus, it had been pointless to wait even before the intention in question was formed; in fact, it seems that the best option would have been to turn on the Laplacean supercomputer immediately after its assembly, find the nearest possible world in which the crime does not take place, and then do whatever is necessary to ensure that the actual world unfolds in the same way.

---

[3]In case anyone were in doubt, the notion of a Laplacean *super*computer is comfortably compatible with determinism. Perfect information with regard to the initial data of the universe and the laws operating upon them plus the capacity to factor in the relevant counterfactual differences is all that the existence of such a device requires, and the last of these requirements seems to me to be obviously agreeable with a deterministic worldview.

This last step could involve various measures, ranging from subjecting the would-be criminal to some educational or psychiatric procedure to incarcerating him or perhaps even preventing his mother from giving birth to him (depending on how difficult he would prove to handle at any given stage of his life). In this connection, it is also worth noting that there might be cases in which waiting for the criminal intention to form is not so much pointless, but fatal: consider the "supervillain" scenario in which the antagonist does not entertain any malevolent plans until he realizes the full extent of his powers, but as soon as this happens, he becomes unstoppable, irreversibly beyond the reach of both prevention and punishment.

In view of the above, it appears wholly arbitrary to claim that preactions can be undertaken in response to certain mental events (e.g., formation of a felonious plan), but not in response to other mental events, which lead to the occurrence of the former (e.g., pure contemplation of one's destructive power), or in response to non-mental events that play a similar role (e.g., biological formation of the future villain's brain). On what grounds can the compatibilist say that only the first item on the above list belongs to the causal chain of events whose unalterable conclusion (given no preactive interference) is the commission of the crime? If, as part of his response, the compatibilist were to contend that the felonious plan is genuinely *willed*, I would say that the same goes for the preceding contemplation of one's destructive power, as well as for all the previous mental events that comprise the causal chain in question, including the criminal's mother's intention to have children.

But even apart from that, I am still unable to understand the sense in which the compatibilist notion of willing can be deemed free. Regardless of whether freedom is to be considered as the counterfactual possibility of being able to have done otherwise or as the defining characteristic of agent-causation (where the characteristics of the agent are not themselves deterministically caused), the compatibilist's account is found lacking.[4] It seems that what he can establish at most is that mental events, as opposed to non-mental events, can offer their subjects an illusion of being the result of

---

[4] A vivid example of such deficiencies are the ones embodied in Daniel Dennett's "evolutionary compatibilism" (Dennett 2003). Dennett claims that free will is implied by the possibility to avoid certain outcomes, and that our evolutionary makeup ensures that this possibility is real (we can dodge bricks etc.). The problem is that he helps himself to a semantic trick and conflates two meanings of "avoidance"—avoiding as in "he avoided the danger" and avoiding as in "this tragedy could have been avoided." Determinism makes no place for the latter: if it is true, then every instance of avoiding something is in fact unavoidable, thus freedom understood as being able to have done otherwise is absent. So is freedom understood in terms of agent-causation, since, on a deterministic worldview, the processes that determine all the agent's choices start long before his birth.

free will, but that can be explained on account of them being conscious, and consciousness does not imply freedom. I believe that the above (admittedly age-old) reservations about the viability of compatibilism are given new force by the observation that undertaking preactions—provided that harm prevention is their purpose—is justified at any point[5] in the causal chain of events whose culmination they are to obviate. Waiting for the supervillain's execution of his criminal plan is as pointless as waiting for the supervillain's realization (or perhaps even formation or maturation) of his powers, since the former is just as determined as the latter. In sum, compatibilism collapses into hard determinism.

But I do not believe that accepting this conclusion should make us hard determinists. Let me therefore outline my final, tentative, libertarian suggestion. The crucial question that needs to be asked in this connection is: could Laplacean calculators really work? If given the exact initial data of the universe, could they really reveal to their users the complete history of cosmic events? This is precisely what I wish to dispute; my claim is that there is nothing nomologically impossible in the notion of an event which is neither determined nor random, an encounter with which would inevitably crash the fantastic devices in question.

Imagine a being (Z) confronted with an apple tree and a pear tree. To make the scenario somewhat simplified, let us suppose that Z possesses two sets of brain cells, 10 brain cells each, responsible for producing a taste for apples and pears respectively. Thus, the intensity of the corresponding food-desires is fully equal. Further, let us assume that the environment (other than the trees) and Z's mental history either do not influence these two desires or influence them to an equal degree. And finally, let us suppose that Z is not wont to establishing a preference in problematic cases by resorting to such procedures as coin flipping. I do not think that any of these assumptions involves a nomological impossibility.

Now, the essential question to ask is: what will Z choose, pears or apples? Can any Laplacean device tell us *that*? It seems to me that it cannot: the chances are even, and the outcome is not determined. But, contrary to what the friends of compatibilism might suggest, it is not random either. If it were to be truly random, then some quantum trigger could cause Z to do virtually anything: recite a poem, do a somersault, climb up and down the tree without picking any fruit etc. Surely, this is not what we should expect. The crux of decision-driven causation is that the range of relevant options is

---

[5] And most justified at the most cost-efficient point.

determined by the data of the environment and the agent's psychology[6] (the agent clearly has to have something to choose *from*), but it is up to him which of these options to pick. I take it to be a fair explication of what Robert Kane called:

> ultimate control—the originative control exercised by agents when it is "up to them" which of a set of possible choices or actions will now occur, and up to no one and nothing else over which the agents themselves do not also have control. [Kane 2003, p. 243]

The final question that determinists would presumably wish to ask at this point is: how do you explain the existence of this mysterious phenomenon? How does it fit the structure of the natural world? My answer is: it does not need an explanation any more than the equally mysterious phenomena of determinism and randomness need one. Some philosophers of mind developed respectable theories according to which consciousness is a fundamental constituent of reality, next to matter and energy (Stapp 1993, Chalmers 1996, Rosenberg 2004). Similarly, I am sympathetic to the view that there are three fundamental modes of causation: deterministic, random and decision-driven. This view seems to be the one that is most compatible with the existence of some of our crucial cognitive skills (e.g., counterfactual reasoning), as well as the only one that can resist the bizarre concept of preactions.

In any event, even if my comparatively short case for ML needs further development, I hope that my elaboration of the problem of preactions will succeed in pointing out that ML is a precondition for PEL, and that compatibilism is a much less tenable position than is usually thought.

References

Beebee, H. 2008. Smilansky's alleged refutation of compatibilism. *Analysis* 68: 258–60.

---

[6]It is a separate question to what extent the agent's psychology is itself determined. I believe that its formation consists of a string of scenarios similar to that considered in the main text: their outcomes are partly determined by hard-wired factors (e.g., genetic makeup) and partly by free decisions as to how to respond to these factors.

Chalmers, D. 1996. *The Conscious Mind*. Oxford: Oxford University Press.

Dennett, D. 2003, *Freedom Evolves*. New York: Viking Press.

Dworkin, R. 2000, *Sovereign Virtue*. Cambridge Mass.: Harvard University Press.

Kane, R., ed. 2003. *Free Will: New Directions for an Ancient Problem*. Oxford: Blackwell.

Kearns, S. 2008. Compatibilism can resist prepunishment: a reply to Smilansky. *Analysis* 68: 250–53.

Rosenberg, G. 2004. *A Place for Consciousness: Probing the Deep Structure of the Natural World*. New York: Oxford University Press.

Smilansky, S. 2007. Determinism and prepunishment: the radical nature of compatibilism. *Analysis* 67: 347–49.

Smilansky, S. 2008a. Prepunishment for compatibilists: a reply to Kearns. *Analysis* 68: 254–57.

———. 2008b. More prepunishment for compatibilists: a reply to Beebee. *Analysis* 68: 260–63.

Stapp, H. 1993. *Mind, Matter and Quantum Mechanics*. Berlin: Springer Verlag.