

Commentary
Living Without Free Will

Derk Pereboom
Cambridge University Press, Cambridge, 2001

Giuliano Torrenco *
giuliano.torrenco@unito.it

The thesis that we utterly lack free will and thus we are not morally responsible for our actions looks difficult to reconcile with the basic features of our ordinary experience. In his influential book, Derk Pereboom argues in favor of it, as the most rational view on agency on the market. In this commentary, I will discuss his view and make few remarks on the connection between certain problems in the philosophy of agency and problems in the metaphysics of time and persistence.

Remember the general framework of the debate over free will and moral responsibility. If someone is morally responsible for her actions, she can deserve praise or blame for them. But moral responsibility requires freedom, since if someone is not free, she cannot deserve praise or blame. The *compatibilist* thinks that being free in the sense required for moral responsibility is compatible with physical determinism. We can be free even if our actions are determined by all what has happened in the past, given the actual laws of nature. The *incompatibilist* disagrees and thinks that if determinism is true, we can't be free in the sense that matters for moral responsibility, and thus we cannot be morally responsible. Traditionally, there are two varieties of incompatibilism. The first is *libertarianism*, who maintains incompatibilism and denies determinism. According to the libertarian, we are free agent in an indeterministic world. Therefore, we are morally responsible. It follows that praise and blame can be rationally justified reactions to human actions. The second one is *hard determinism*. According to the hard determinist, incompatibilism and determinism are both true. Thus, we are agents who lack free will and who inhabit a deterministic world. We lack moral

* University of Turin

responsibility and judgments of moral praise or blame are always irrational – they only seem to be justified.

Pereboom defends a form of incompatibilism – which he calls *hard incompatibilism* – that is neither libertarianism nor hard determinism, although it is clearly closer to the latter than to the former. This is true not only because hard determinism is subsumed under hard incompatibilism: being a hard determinist is a way of being a hard incompatibilist, although not always the other way around. But more profoundly because Pereboom's incompatibilism, in a sense, embodies the gist of hard determinism, without entailing determinism across the board. The idea is that neither deterministic aspects of our agency nor indeterministic ones are compatible with the tenet that we are morally responsible and free. Not all incompatibilist positions that refute both the libertarian notion of free will and determinism share the notion of agency that Pereboom outlines. The *no-free-will-either-way* position, for instance, maintains that we cannot be free and morally responsible, regardless of whether determinism or indeterminism is true. Galen Strawson defends one version of this view. According to Strawson, it is metaphysically impossible for human beings like us to be free agents. That is why, no matter whether determinism is true or not, we cannot be morally responsible. According to a even stronger version, free will is not just metaphysically impossible, but even conceptually so: the very notion is contradictory.

Contrary to those positions, Pereboom argues that the concept of free agency is both logically and metaphysically well-behaved. Besides, the difference between living in a deterministic world and living in an indeterministic world is relevant for human morality and the rationality of our moral appraisals. Between the two main varieties of libertarian incompatibilism – *leeway* incompatibilism and *causal history* incompatibilism – he thinks that only the second one catches what is essential to freedom and responsibility. Leeway incompatibilism rests on the tenet that having alternative possibilities is not only a necessary condition for freedom and responsibility, but it is a condition with an explanatory import. We are free, morally responsible and thus blameworthy or praiseworthy for our deeds *because* we could have done otherwise than we actually did. Whereas according to causal history incompatibilism, having alternate possibilities as such – even if it turned out to be a necessary condition for being free – cannot explain human freedom and responsibility. An action is free not because the agent could have done otherwise, but because it is not produced by a process that traces back to causal

factors beyond the agent's control, as deterministic processes typically are. In other words, what accounts for an agent being free and morally responsible for her actions is the causal history through which the agent has arrived to those actions. But it is not indeterminism *per se* what makes an agent free. We can think about that in terms of the distinction between open and close future. There is a sense of "open" in which if the future is open, then there is nothing in the present that settles what will be the case tomorrow. Suppose that we live in a universe in which the future is open in that sense: are we thereby free? No, because many things in the future may be unsettled while our decisions being inescapable consequences of factors beyond our control. Thus indeterminism is not a necessary condition for being free.

However, if we share the incompatibilist intuition that if everything is already settled in advance than we cannot be free, we may think that a open future in the indeterminist sense is at least necessary condition for being free. Frankfurt-style examples, though, provide powerful objections to the idea that someone is free and responsible only if she could have done otherwise than she actually did. Suppose that George has to decide whether to lie about his taxes or be honest. Eventually, he decides to act immorally and to lie. However, unbeknownst to him, a neuroscientist has implanted in his brain a device that can detect his intentions. The device is such that were he to form the intention to act honestly with respect to fiscal behavior, it would intervene and make him act immorally instead. In such a situation, George could not have done otherwise than he actually did, and thus the leeway condition for being free is not fulfilled. Yet, we share the intuition that, in the situation described, since the device implanted by the neuroscientist has not intervened, his action has been freely chosen. Contrary to what the leeway incompatibilist theory predicts, he is morally responsible and blameworthy for what he has done even though he could have not done otherwise. If those counterexamples go through, having alternative possibilities is not a necessary condition for being free and the whole project of leeway indeterminism fails. Note that Frankfurt-style counterexamples do not impinge on the requirement of a indeterministic condition on the causal history of the decision of the agent. Indeed, in all such cases, the lack of alternative possibilities is a consequence of "external" factors concerning the situation in which the agent find himself. If we changed the story and assume that the device influenced the causal history in a way to make it deterministic, or in other ways beyond the agent's control, our intuition that an agent is free and morally accountable would fade away.

However, classical Frankfurt-style examples cannot force the libertarian to a causal history version of the position. The intuition that having the capacity of doing otherwise is relevant for our freedom and responsibility is hard to give up. The libertarian can still argue that the intervening device in Frankfurt-style's examples does not touch the agent's capacity of doing otherwise that is essential for moral responsibility. It is true that the agent cannot *act* differently, because if he decided to do so, the device would intervene and prevent him from acting differently. However, an alternative possibility condition is still at work in explaining why even in those circumstances the agent is free. If the agent could not *decide* to do otherwise, we would not have the intuition that she acted freely and she is morally responsible for her action. As it is sometimes put, the possibility of a "flicker of freedom" is required for moral responsibility. Does this fact confirm that the fundamental idea of leeway incompatibilism, i.e., that possessing alternative possibilities is a fundamental factor in explaining freedom and responsibility? No, Pereboom seems rather to think that if the incompatibilist takes seriously the challenge set by Frankfurt-style cases, an indeterminist condition on the causal history of the agent's actions will emerge. Even if eventually the leeway incompatibilist may be right in claiming that there is an alternative possibilities condition necessary for freedom, that is so only in virtue of the holding of a condition on the causal history of the agent's decision.

Consider the compatibilist objection to the tenet that the possibility that a flicker of freedom occurs is necessary for our responsibility. There are more complex Frankfurt-style cases, in which the device implanted by the neuroscientist is able to detect some previous sign of the agent's decision to act immorally. Thus, if the sign does not manifest itself, the device intervenes and forces the agent to decide to act immorally. Suppose that the sign is blushing at a certain moment t , and that it *does* manifest itself at t . Of course, if indeterminism is true, it is still true that the agent could have not blushed at t , but this flicker is not "robust" enough to ground freedom in the sense required for moral responsibility. Indeed, we have the intuition that whether the agent blushes or not at t is irrelevant for the moral import of her action. It cannot be the occurrence of the blushing or the lack thereof *per se* that accounts for the agent's moral responsibility. The alternative possibilities that can justify the leeway incompatibilist condition on freedom must be such that whether the agent goes for one or the other is relevant for her moral responsibility. If she is blameworthy, then it has to be the case that had she done otherwise, she would

have been praiseworthy and vice versa. But clearly she is not blameworthy for blushing at t as such. Whether the agent decide or not to act immorally would be a flicker robust enough to ground moral responsibility, but in the new scenario it is not the case that the agent could have decided otherwise, and thus it is not a flicker at all.

Pereboom sides with the libertarian in the debate over this refined Frankfurt-style cases, since he thinks that although it is not the presence of alternative possibilities as such what makes the agent free and responsible, the refined version of Frankfurt-style arguments still leave the leeway core intuition intact. It is important to notice that the link between the sign and the decision has not to be either deterministic or in any way sufficient for causally determining the agent's decision. If that were the case, then the intuition that the agent is free will be too weak to survive the incompatibilist standards. If the agent is forced to act in a certain way by the occurrence of the sign, then he is not free in any interesting sense. But if the sign is not sufficient to determine that the agent will act in a certain way, than it is still in the power of the agent to do otherwise, and the example does not disprove the alternative possibilities condition on freedom. Indeed, the agent is free only if the possibility of a robust flicker of freedom is still open to her. And the leeway incompatibilist can claim that her moral responsibility is explained precisely by the fact that she is praiseworthy or blameworthy depending on which way she goes.

However, Pereboom parts company with the libertarian because he does think that there are Frankfurt-style cases in which the leeway compatibilism fails. Suppose the sign for deciding to act immorally at t' is that the agent at t does not consider some strong moral reason to act morally, and suppose that considering such reason is not causally sufficient for her decision to act morally, but only necessary. Furthermore, the neuroscientist has implanted a device that is idle in so far as it does not detect any activity of considering moral reasons at t , but it forces the agent to act immorally at t' in case it detects moral considerations at t . Now, the agent at t does not engage in any moral considerations, and consequently the device does not intervene. We have the intuition that the agent is free, even if it is not the case that the agent could have done otherwise, not even in the sense that she could have decided otherwise. Letting apart the details of the discussion, which have given raise to much interest in literature, what is relevant here is to stress that Pereboom version of the Frankfurt-style objection to the leeway incompatibilist is designed to show that the causal history condition is fundamental for

explaining moral responsibility. After all, why in Frankfurt-like cases we have the intuition that the agent is still free and morally responsible? Insofar as the causal history of the decision is not touched by the intervening device or other factors that are not in power of the agent, the intuition of freedom is left untouched by the presence of intervening devices – no matter how subtle and “invasive” they are. Note, for instance, that if we make the connection between the sign and the decision too strict in terms of sufficient causal determination, then we lose the intuition that it is still in control of the agent to do otherwise. That is because if the link is causally determining, then the causal history of the decision would contain aspects that are beyond the agent’s control. Therefore, the relevant condition for having freedom and responsibility – the conditions with explanatory power – is having an indeterministic causal history, such that allow for the agent to have control over her decision. If this condition holds, then we can derive some sort of alternative possibilities conditions too – but the core of the notion of freedom that is relevant for moral responsibility does not lie in the presence of alternative possibilities, rather in having control over one’s own decisions.

If the condition of alternate possibilities does not catch the core of our notion of freedom and moral responsibility, the problem of determining whether someone “could have done otherwise” is no longer crucial for establishing moral responsibility and freedom. And this is good news because the debate on freedom and moral agency risks to wind up in a stalemate by focusing on the he proper analysis of “could”. A compatibilist would argue that a counterfactual analysis of “could” is required in such cases. Very roughly, if there are possible worlds close enough to ours in which the agent acts otherwise, then it is true in the actual world that she could have done otherwise. But the fact that her choice is causally determined by previous facts beyond her control does not imply that its occurrence is metaphysically necessary, i.e., that there are no possible worlds in which she acts otherwise. Thus, since the agent could have done otherwise, she is free even if she inhabits a deterministic world. The incompatibilist objects to a counterfactual analysis of “could” here, and argue that there is a sense in which if determinism is true, then the agent could have *not* done otherwise, and thus it is not free in a deterministic world. Which sense of “could” should we consider here? If we maintain that the alternative possibility condition is explanatory central for freedom, both are relevant for the question whether the agent is free or not.

And – most importantly – both senses are legitimate, since the nicely match the compatibilist and the incompatibilist supporting intuitions respectively.

The version of incompatibilism that Pereboom puts forward is immune to the risk of finding itself in such a dead-end. According to causal history incompatibilism, we are free only if the causal history of our choices and deliberations involves some essential factor that is under our control. If an agent is morally responsible for her decision to perform a certain action, then the production of this decision must be something over which the agent has control, and an agent is not morally responsible for the decision if it is produced by a source over which she has not control. The point is not only that if an action is a inevitable consequence of what has happened so far in the universe (given the actual laws of nature), then it cannot be a free action and the agent cannot be morally responsible for it (note that leeway incompatibilism, too, may be claimed to catch this aspect). The point is that we are justified to believe that the agent is not free and morally responsible for her actions only in case that her actions have originated from something over which the agent has not control. But being a deterministic consequence of previous events is not the only way in which an event can escape our control. Also a event that happens for no cause at all or randomly may be completely beyond our control.

It is crucial to stress here that indeterminism as such is no warrant of freedom and responsibility – as compatibilists have often stressed. And for the same reason that leeway incompatibilism fails to catch the core of the notion of freedom. Suppose that our world is indeterministic, and more precisely, the processes through which a agent gets to a decision are indeterministic. What we decide will be a consequence of which ones among the alternative possibilities have turned out to be actual, and if we have no control over those events, then we are not free. What, as a matter of fact, is beyond our control can account for our freedom no more if it is a consequence of a indeterministic process than if it is the outcome of a deterministic process. Therefore, no matter whether the causal history of our decisions is deterministic or indeterministic, if there are no crucial elements of it that are under our control, we cannot be free. However, can the processes underpinning our deliberations be such that they are in some relevant way under our control? Libertarians think it can, whereas hard indeterminists maintain that all our decisions are determined causally by things outside our control, since they are *alien determinist events*. Pereboom sides with the libertarian in maintaining that it is

metaphysically possible that a human agent be free because in control of her deliberations. However, he sides with the hard determinist in maintaining that we do not have free will and moral responsibility. His reasons for that claim, however, are not grounded in the truth of determinism: rather, there are very good empirical reasons to believe that we are not in control of any of the events that constitute our decisions.

Actually, with respect to what he calls “event-cause libertarianism” as opposed to “agent-cause libertarianism”, Pereboom’s position is slightly stronger, since he thinks that event-cause libertarianism encompasses a notion of agent such that, at least by metaphysical necessity, is not in control of her deliberations. I think that the idea of event-cause libertarianism can be made more precise by appealing to the underlying metaphysics of persistence, and in particular, to the distinction between endurantism and perdurantism. According to the perdurantist, agents – as any other entity that persists in time, namely that exists at more than one time – are nothing over and above mereological sums of instantaneous events. Those events are the temporal parts of the agents – those commonly said to be the phases of the agent’s life. Within this framework, it is easy to tell what is an agent’s decision: it is a temporal part of the agent. What causes a decision, though? If it is the outcome of a deterministic process, it is caused by former parts of the agent in such a way that the agent has no control over the process and thus she is not free. If it is a *truly random event*, viz. something that happens with no cause whatsoever, the agent will not have control over it either. But even if it is a *partially random event*, which the agent cannot causally determine, the agent will not have enough control over her decisions to be free. Therefore, if decisions are events either without a cause or caused by other events, as in the event-based version of libertarianism, it is hard to see how there can be decisions over which the agent has enough control to be free and morally responsible. Since, even the non-random part of the determination can only be another event over which the agent has no control. And the same goes if the libertarian insists that the causally determining factors are things like the agent’s character or her capacities. In the event-based version of the theory, what causes an agent to have her actual character cannot be something over which the agent has control.

However, there is a way to add the kind of control required for freedom to the indeterminist picture of the libertarian. In so far as it is coherent to maintain that the agent herself, and not an event (even if one strictly connected

to the agent, as one of her temporal parts), causes her decisions without being determined by factors beyond her control, the notion of a free agent in the sense required for moral responsibility and thus the notion that the libertarian needs to state her position is coherent. Agent-cause libertarianism is precisely the view that it is a primitive feature of the agent to be such a causal source of her actions. Although we can make sense of the idea that the agent as the whole composed by temporal parts, and not any of the parts as such, is the ultimate free cause of her actions, I think that an endurantist metaphysics makes the picture far neater. According to the endurantist, the agent – as any entity that persists in time – persists by being wholly present at each moment of her existence. That is, it is the agent itself, and not any of her temporal part that we find in each phase of her life. Within the framework of an endurantist metaphysics it is clear how the causal relation underpinning the agent's choices looks like: one of the terms of the relation is the agent, the other term is an event, namely the choice that the agent has caused to occur (or to whose occurrence the agent has contributed fundamentally).

Now, endurantism is a less revisionary metaphysics than perdurantism, i.e., it is closer to common sense. However, it is also a position less sympathetic to hard sciences. This is true in general, but it is even more apparent in the present case. Modern science makes the notion of an agent as the free cause of her decisions suspicious. The non-reductive materialist strategy to accommodate agent causation within the physical world looks the most attractive, but – as Pereboom convincingly argues for a whole chapter – is not better off than the alternative strategies. At the end of the day, the notion of agent causation of the libertarian violates well-established scientific conceptions. Therefore, even if freedom and moral responsibility are coherent notions, and it is metaphysically possible for a free agent to exist, we are not likely to live in a world inhabited by free agents, and thus we are not justified in seeing us or the other as morally blameworthy or praiseworthy. In other words, the best libertarian version of causal history incompatibilism, namely the agent-based one, has to be abandoned, and the only plausible incompatibilist alternative left is hard incompatibilism. The conclusion is that we are not in control of any of the outcome of our choices, because the causal history of our decisions is entirely made of events over which we do not have control: alien deterministic events, truly random events or partially random events.

Hard incompatibilism has to be defended also from the compatibilist challenge. A compatibilist could agree with Pereboom's picture of an agent as causally determined both in the deterministic and in the indeterministic aspects of the processes underpinning her decisions. Yet the compatibilist ascribes moral responsibility and freedom to humans. According to compatibilists, indeed, causal determination – the sort of lack of control that Pereboom ascribes to agents in ordinary cases – does not exclude free will and grounded ascriptions of responsibility. Therefore, Pereboom has to distinguish his position from its compatibilist counterpart, namely the position embracing both causal determination and moral responsibility, and to defend it as the only viable alternative. To that effect, Pereboom argues that compatibilism fails to spot the relevant similarity between ordinary decisions in which the agent is causally determined by factors behind her control and situations in which the decision is the outcome of a covert manipulation. Since our intuitions in cases of covert manipulations are that the agent is not free and responsible, we should conclude that causal determination in ordinary situations, too, is incompatible with the assumption that the agent is free and responsible. In the present context, I will not discuss the “four-case argument” that Pereboom puts forward to defend his tenet that the two situations are similar in the relevant respect and thus compatibilism fails. Rather, I wish to focus on the modal status of the incompatibilist's notion of determination.

Although Pereboom is right in claiming that the focus on the issue of the proper analysis of “could” leads the debate on free will to a stalemate, I do not think that that is true with respect to all modal considerations about decisions and actions, in particular with respect to the distinction between determination of the future and necessity of the future. Firstly, it is crucial for hard incompatibilism that causal determination does not imply metaphysical necessity. If every human choice is metaphysically necessary, then its occurrence is entailed by the state of the world up to the moment of its occurrence together with the laws of nature (since anything entails a necessary truth), and hard indeterminism collapses on hard determinism. Secondly, the hard incompatibilist's notion of causal determination should not imply nomological necessity either, and essentially for the same reason: if the choices of the agent are determined by what happened in the past together with the laws of nature, indeterminism cannot be true, and full-fledged hard incompatibilism follows the same fate. Can the hard incompatibilist resort to a notion of “logical” determination, which is weak enough to allow for both

determinist and indeterminist factors to enter the causal history of the agent's decisions? If "logical determination" here means simply that every statement about an agent's choice is bivalent (either true or false, but not both or neither) then I maintain that he can. Maybe there is a stronger notion of indeterminism to the effect that statements concerning our choices are logically undetermined, but physical indeterminism (the tenet that the history of the world up to a given moment together with the laws of physics does not settle all aspects of the future) surely is compatible with the claim that statements concerning our choices are logically determined. Moreover, once it's clear that the hard determinist *should* endorse this notion of determination, because it is the only plausible alternative left, it becomes also clear that in the hard incompatibilist's picture, it is *not* the case that the agent lacks freedom *in virtue* of the outcome of her choices being logically determined. And this is an important difference between hard incompatibilism and hard determinism, which can rest on the idea that we lack moral responsibility because every action of ours is determined in advance. To see the point, consider the following: even in an indeterministic world *with free agents*, the logical sense of determination can be maintained (for instance, if something like "the thin red line view" is true). Therefore, if the agent lacked freedom only in virtue of her choices being logically determined, hard incompatibilist would overgeneralize to libertarianism and it would turn out to be incoherent. That situation forces the hard incompatibilist to put the crucial distinction between the libertarian and himself in some other feature of the causal processes that leads to decisions. For the libertarian there is a causal link between the agent and the choice, which is under the agent's control, whereas for the hard incompatibilist there is no causal link of this sort, because all causal relations that underpin the agent's decisions are relation between event.

In the last three chapters, Pereboom focuses on the moral consequences of his theory, in order to defend it against the charge of making morality impossible. The central idea here is that moral responsibility is only an aspect of morality, and it is not even likely to be the most essential one. Indeed, praise and blame seem to be relevant only to the "irrational" part of morality, and moral value is untouched by them. Actions can be either morally good or wrong, even if they are never praiseworthy or blameworthy. Pereboom does not deny that sometimes an emotional twist may be beneficial for morality, but he argues that for the most crucial aspects of morality there are incompatibilist "surrogates" for praise and blame. In any case, the overall appraisal of the

moral role of rationally ungrounded emotions seems to suggest that we should dispense with them. Pereboom, thus, does not think – as certain hard determinists do – that for pragmatic reasons we should act *as if* we were morally responsible. Living by thinking that we lack freedom and responsibility is morally desirable and within our ordinary capacities. Even if the argument of these chapters probably will not convince anyone who, at this point, is not already both incompatibilist and non-libertarian, they provide lively challenges for libertarians and complete the overall plausibility of the hard incompatibilist view of reality.