

STUDI

Response to Daniel Dennett on Free Will Skepticism

Derk Pereboom^(α)

Ricevuto: 17 febbraio 2017; accettato: 24 agosto 2017

Abstract What is at stake in the debate between those, such as Sam Harris and me, who contend that we would lack free will on the supposition that we are causally determined agents, and those that defend the claim that we might then retain free will, such as Daniel Dennett? I agree with Dennett that on the supposition of causal determination there would be robust ways in which we could shape, control, and cause our actions. But I deny that on this supposition we would have the control in action required for us to basically deserve to be blamed, praised, punished or rewarded. In this response, I argue that this is the core issue that divides compatibilists and incompatibilists about free will and causal determination, and that the incompatibilist position is the right one to accept.

KEYWORDS: Sam Harris; Daniel Dennett; Free Will Skepticism; Compatibilism; Incompatibilism

Riassunto *Risposta a Daniel Dennett sullo scetticismo circa il libero arbitrio* – Qual è la posta in gioco nel dibattito che vede contrapporsi chi – come Sam Harris e me – sostiene che non avremmo libertà di volere sulla scorta dell'ipotesi per cui siamo agenti causalmente determinati e chi, al contrario – come Daniel Dennett – difende l'idea che possa darsi un libero volere? Concordo con Dennett circa il fatto che, anche nell'ipotesi della determinazione causale, resterebbe lo spazio per sostenere che per vari e importanti aspetti saremmo comunque noi a modellare, controllare e causare le nostre azioni. E tuttavia rifiuto che su questa base avremmo il controllo in azioni richieste per meritare di essere biasimati, lodati, puniti o premiati. A mio avviso questo è l'elemento cardine che divide i compatibilisti dagli incompatibilisti all'interno del dibattito sul libero arbitrio e sulla determinazione causale e ritengo che la posizione corretta sia quella incompatibilista.

PAROLE CHIAVE: Sam Harris; Daniel Dennett; Scetticismo verso il libero arbitrio; Compatibilismo; Incompatibilismo



IN THE FREE WILL DEBATE, I'm classified, like Sam Harris, as a free will denier and an incompatibilist. But these classifications obscure important distinctions. I'm in many respects in agreement with Daniel Dennett's position, and he self-identifies as a compatibilist. I agree, for example, that on the supposition of causal determination there are ro-

bust ways in which we shape, control, and cause our actions. The terminology in the free will debate is at times too coarse-grained, and participants on occasion talk past each other as a result. But this problem can be avoided if we take care to be clear about what we mean.

The terms "free will" and "moral respon-

^(α)Sage School of Philosophy, Cornell University, Goldwin Smith Hall 218 – 14853-3201 Ithaca, New York USA

E-mail: dp346@cornell.edu (✉)



sibility” are used with different senses, and the type of free will or control required for moral responsibility given several of these senses is uncontroversially compatible with the causal determination of action by factors beyond our control. I take “free will” to refer to the strongest sort of control in action required for a core sense of moral responsibility.¹ That sense of moral responsibility, the one at issue in the free will debate, is set apart by the notion of *desert*.² For an agent to be morally responsible for an action in this sense is for it to be hers in such a way that she would deserve to be blamed if she understood that it was morally wrong, and she would deserve to be praised if she understood that it was morally exemplary.

It may be that the desert at issue is *basic* in the sense that the agent would deserve to be blamed or praised just because she has performed the action, given an understanding of its moral status, and not, for example, merely by virtue of consequentialist or contractualist considerations. A belief that an agent is morally responsible in this sense at least typically accompanies expressions of reactive attitudes such as moral resentment and indignation, and it is thus closely related to the notion of moral responsibility that P.F. Strawson invokes.³ But there are other senses of moral responsibility that are not at issue in the free will debate, and are not challenged by skeptical arguments about free will, and the legitimacy of some of them is an important feature of my position.

Some philosophers identify themselves as compatibilists because they hold that some non-basic-desert notion of moral responsibility, often one they regard as sufficient for the moral life, is compatible with determinism.⁴ But if “compatibilism” is defined so that such a position turns out to be compatibilist, virtually everyone in the debate stands to be a compatibilist. Frank Jackson says:

What compatibilist arguments show, or so it seems to me, is not that free action as understood by the folk is compatible with

determinism, but that free action on a conception near enough to the folk’s to be regarded as a natural extension of it, and which does the theoretical job we folk give the concept of free action in adjudicating questions of moral responsibility and punishment, and in governing our attitudes to the actions of those around us, is compatible with determinism.⁵

Dennett specifies that his compatibilist notion of free will can «play all of the valuable roles free will has been traditionally invoked to play».⁶ However, Stephen Morris argues that Dennett

has defined the concepts of “free will” and “moral responsibility” in such a way as to eliminate any substantive difference between the “compatibilist” position he defends and the hard determinist position that philosophers typically understand as being substantively different from compatibilism.⁷

The same might be said for Jackson’s view, depending on the details. Given Jackson’s and Dennett’s characterizations, virtually everyone would qualify as a compatibilist. I’d be one, for example, and it’s likely that Sam Harris would be one as well. If this is the notion of compatibilism that those in the Chalmers and Bourget survey had in mind, it would be surprising that as many as 12% of those surveyed are in fact incompatibilist free will deniers.⁸ I’ve been working in this field for a quarter century, and no clear example comes to mind.

Core disagreements with Dennett

Let me now set out the issue on which Dennett and I differ, and my sense is that here Harris and I concur in disagreeing with him. The manipulation argument against compatibilism brings this disagreement to the fore. The core idea of the manipulation argument is that an action’s being produced by a deterministic process that traces back to

factors beyond an agent's control, even when it satisfies all the conditions on moral responsibility specified by the contending compatibilist theories, presents no less of a challenge to basic desert responsibility than does deterministic manipulation by other agents. My multiple-case manipulation argument⁹ sets out several manipulation examples, the first of which features the most radical sort of manipulation consistent with the proposed compatibilist conditions. The subsequent cases are progressively more like a final example, which the compatibilist might envision to be ordinary and realistic, in which the action is causally determined in a natural way. A challenge for the compatibilist is to point out a relevant and principled difference between any two adjacent cases that would show why the agent might be morally responsible in the later example but not in the earlier one.

We don't have to examine the entire argument to see where the disagreement with Dennett lies. The first manipulation case from my 2014 book will do. The situation features Plum, who murders White for reasons of self-interest:

Case 1: A team of neuroscientists has the ability to manipulate Plum's neural states at any time by radio-like technology. In this particular case, they do so by pressing a button just before he begins to reason about his situation, which they know will produce in him a neural state that realizes a strongly egoistic reasoning process, which the neuroscientists know will deterministically result in his decision to kill White. Plum would not have killed White had the neuroscientists not intervened, since his reasoning would then not have been sufficiently egoistic to produce this decision. But at the same time, Plum's effective first-order desire to kill White conforms to his second-order desires. In addition, his process of deliberation from which the decision results is reasons-responsive; in particular, this type of process would have resulted in Plum's re-

fraining from deciding to kill White in certain situations in which his reasons were different. His reasoning is consistent with his character because it is frequently egoistic and sometimes strongly so. Still, it is not in general exclusively egoistic, because he sometimes successfully regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. Plum is also not constrained to act as he does, for he does not act because of an irresistible desire – the neuroscientists do not induce a desire of this sort.¹⁰

Does Plum deserve to be blamed or punished for what he's done? I don't think so, and many others surveyed. If we asked whether Plum basically deserves to be blamed or punished, perhaps even more would agree with me. True, there are those who hold that Plum does deserve to be blamed and punished.¹¹ But there's plenty of opposition to this compatibilist line of thought.

Another question one might pose is: Could Plum have done otherwise given that the compatibilist conditions are met? We can all agree that he has the general sort of ability to do otherwise that Dennett spells out. He indeed does have the general ability to kill and to refrain from killing. But we can ask: could he have exercised his general ability to refrain from killing on this particular occasion?¹² And now a common answer may be "no." Further, suppose that on this occasion he could not have exercised his general ability to refrain from killing. We can then ask: Does he deserve, or basically deserve, to be blamed or punished? Here again, a common answer might well be "no."

■ A notion of moral responsibility that's good enough

Our practice of holding each other morally responsible is complex. It involves a number of different aims, and a range of responses justified by those aims. In recent years a number of theorists have argued that this complexity

can be unified, that there is ultimately a single notion of moral responsibility that unifies the practice. I believe that a view of this sort misrepresents the practice to at least some degree, and here my potential allies include Gary Watson, Dana Nelkin, and David Shoemaker.¹³ With them, I contend that a certain kind of pluralism about the practice is true.

It's mistaken to claim that the term "moral responsibility" has a single sense. Instead, there are various senses of "moral responsibility" corresponding to multiple aims and justifications, all of which are aspects of our general practice of holding morally responsible. Purely linguistically, this is credible. Often linguistic terms that have a long history of use have a number of senses. Even if a referring term with such a history originally had just one specific referent, over time it is apt to be applied to similar but different referents, acquiring different senses as a result. The terms "free will" and "moral responsibility" plausibly have this feature.

It's clear that the usual holding-accountable aspect of the practice features the notion of desert. There may in addition be senses of moral responsibility that involve a non-basic variety of desert. Essentially forward-looking notions of holding agents deserving of blame and punishment have been defended on consequentialist or contractualist grounds.¹⁴ But there might be no such thing, and as, we've seen, the manipulation argument provides an indication that any notion of desert is incompatible with causal determination. On one type of revisionary account, our practice of holding agents morally responsible in a desert sense should be retained, not because we are in fact morally responsible in this sense, but because doing would have the best consequences relative to alternative practices. Dennett¹⁵ advocates a version of this position, as does Manuel Vargas.¹⁶ One question for Dennett and Vargas is whether, given their positions, what we call deserved responses are really just negative or positive incentives. We might not ordinarily regard them merely as incentives, but on this view their justification is ultimately forward-looking,

founded solely on the value of consequences that we expect to result. And if these responses are in fact just incentives, then it would seem that they are not also deserved.

There are reasons to be skeptical of any notion of moral responsibility that involves desert even independently of any considerations regarding free will. One worry for the basic desert sense is that for an agent to basically deserve a harmful response she must have a kind of free will that is unavailable to us, and the free will skeptic contends that this concern can't be successfully countered.¹⁷ As noted, one might argue that some desert sense of moral responsibility can or should be retained because doing so stands to bring about good results, but on this view what might appear to be deserved responses would seem to be mere incentives. Another worry is that for a number of contending general normative ethical theories the notion of desert seems to have the role of an awkward supplement. Any place for desert in typical consequentialist views is uncomfortable, and despite Kant's invocation of desert in justifying criminal punishment,¹⁸ that appeal appears not to be justified by any formulation of the Categorical Imperative, which he held to be the supreme and comprehensive moral principle. A further issue is that negative desert, at least in its basic form, would seem to involve the idea of harm as an intrinsic good, which is dubious.

Motivated by these concerns, I've proposed a view that rejects all desert-involving senses of moral responsibility. The notion of moral responsibility, and blame in particular, that I develop and endorse¹⁹ is largely forward-looking. Blame is, in its paradigm cases, a kind of calling to account, and is justified by these forward-looking elements:

- (1) The right of those wronged or threatened by wrongdoing to protect themselves and to be protected from immoral behavior and its consequences.
- (2) The good of reconciliation with the wrongdoer.

- (3) The good of the moral formation of the wrongdoer.

Immoral actions are often harmful, and we have a right to protect ourselves and others from those who are disposed to behave harmfully. Immoral actions can also impair relationships, and we have a moral interest in undoing such impairment through reconciliation. And because we value morally good character and resulting action, we have a stake in the formation of moral character when it is beset by dispositions to misconduct.

There is an account of praise that corresponds to this conception of blame. Of the three goals of blaming, moral formation, protection, and reconciliation, the one most clearly amenable to praise is moral formation. We praise an agent for a morally exemplary action to strengthen the disposition that produced it. This can have a protective function, since strengthening such dispositions has the effect of reducing the incidence of dangerous behavior. Corresponding to reconciliation is the notion of celebrating successes in a relationship. Praising an action has this celebratory function as well.

What is morally distinctive about blaming given the rejection of desert? I propose that the notion of *moral protest* is the essence of blame in the sense I'm setting out.²⁰ The inspiration comes from Pamela Hieronymi, who argues that resentment is best understood as moral protest: <resentment protests a past action that persists as a present threat>.²¹ Resentment is not a feature of the sense of blame I have in mind. But on Hieronymi's account a key function of blaming someone is to protest a past action of his that persists as a present threat, and I accept, in the account of blame I endorse, that this is one of its core functions. Angela Smith's recent theory of blame also features a notion of moral protest, this time as the essence of blame.²²

I endorse the following account:

Moral Protest Account of Blame: For B to blame A is for B to issue a moral

protest against A for what B represents (either truly or falsely) as A's immoral conduct.

The immoral conduct will typically be an immoral action, but there are examples in which the action considered separately from the reasons for which it's performed is not wrong, but the reasons make the overall conduct wrong.²³ At times blame is misplaced, since no immoral conduct has taken place, but the protest can still count as blame. This can happen when X believes Y to have acted badly but the belief is false, due perhaps to misinformation or improper consideration of evidence. This can also occur when Y does not believe that X acted badly but nonetheless represents X as having acted immorally, as in cases of politically motivated false accusation. It's often the case that blame functions, as in Hieronymi's account, as a moral protest of a past action that persists as a present threat, and I agree that this is one highly important objective for blame. But not all blame has this point, as when we blame the dead, or blame someone who is alive but lacks a persisting disposition to act badly – someone, for instance, who has already undergone moral reform. In such cases protest can yet have the function of explicitly noting immoral conduct as immoral, which might also have the aim of general moral improvement. In the example of the already-reformed wrongdoer, blame might still function as a step in the process of reconciliation. While this moral protest account captures much of blaming behavior, it is consistent with the revisionary view that desert presuppositions, which are widespread, have no justified role in blaming, and that its legitimate aims are instead those I've set out.²⁴

■ A way to deal with criminal behavior that's good enough

We also don't need the notion of desert for a justifiable and workable way to deal with criminal behavior. I argue²⁵ that it is not

the practice of holding criminals responsible at all that's undermined by scepticism about free will, but only the justification of criminal punishment that depends on desert. In particular, it's the retributivist justification of punishment, according to which the punishment of a criminal is justified on the ground that he deserves it just because he has knowingly committed a serious offense, that would be undermined. By contrast with retributivism, justifying punishment on the ground of general deterrence does not seem to conflict with the denial of free will as the control in action required for desert responsibility. I agree that punishment can be justified on the ground of general deterrence as long as the punishment is not so severe that it undercuts a reasonable level of flourishing. Otherwise, the concern is that the person is being used as mere means to the safety of others. In my view it's the right to life, liberty, and physical security of the person that has the key role in the use objection to general deterrence. There is a heavily weighted presumption against punishment as use, where that involves intentional killing, confining, or infliction of severe physical or psychological harm. However, consider monetary fines, when they don't hinder survival or the living at a reasonable level of flourishing. I've argued that the general deterrence involving monetary fines may be in the clear.²⁶ It's plausible that the same can hold for short prison terms.²⁷

But what about dangerous crime that is not deterred by monetary fines or short prison sentences? I've developed the view²⁸ that there is a justification for prevention of dangerous crime that is neither undercut by free will scepticism nor by other moral considerations. This theory is based on the right to self defense and defense of others, and draws an analogy between treatment of dangerous criminals and treatment of carriers of dangerous diseases. Those infected with such diseases are not responsible in this (or in any) sense for the danger they pose, but we generally agree that it is sometimes justifiable to quarantine them nevertheless. But then, even

if a dangerous criminal is not morally responsible for his crimes in the basic desert sense, it would be as legitimate to preventatively detain him as to quarantine the non-responsible carrier of a serious communicable disease. Threats of less dangerous crimes would justify only milder forms of incapacitation than preventative detention, for the reason that the right of self-defense and defense of others justifies only the minimum harm that protection requires. I contend that the resulting account demands a degree of concern for the rehabilitation and flourishing of the criminal that would alter much of current practice. For just as fairness counsels that we attempt to cure the diseased we quarantine, so fairness would recommend that we try to rehabilitate the criminals we detain or otherwise incapacitate.

Thus despite the many similarities between Dennett's view and mine, there is a key difference. I deny that causal determination is compatible with the control in action required for moral responsibility in any desert sense, and in the basic desert sense in particular. I believe that this is no small matter, since relinquishing desert would stand to significantly alter our practice of holding responsible and our ways of dealing with criminal behavior.

Notes

¹ A. MELE, *Free Will and Luck*, Oxford University Press, Oxford 2006, p. 17.

² See J. FEINBERG, *Doing and Deserving*, Princeton University Press, Princeton (NJ) 1970; D. PEREBOOM, *Living without Free Will*, Cambridge University Press, Cambridge 2001; D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, Oxford University Press, Oxford 2014; J. FISCHER, R. KANE, D. PEREBOOM, M. VARGAS, *Four Views on Free Will*, Blackwell, Oxford 2007; T.M. SCANLON, *Giving Desert Its Due*, in: «Philosophical Explorations», vol. XVI, n. 2, 2013, pp. 101-116.

³ See P.F. STRAWSON, *Freedom and Resentment*, in: «Proceedings of the British Academy», vol. XLVII, 1962, pp. 187-211.

⁴ See F. JACKSON, *From Metaphysics to Ethics*, Oxford University Press, Oxford 1998; D.C. DEN-

NETT, *Freedom Evolves*, Viking, New York 2003, p. 225. See also M. VARGAS, *Revisionism and Response to Fischer, Kane, and Pereboom*, in: J. FISCHER, R. KANE, D. PEREBOOM, M. VARGAS, *Four Views on Free Will*, cit., pp. 126-165 and pp. 204-219.

⁵ F. JACKSON, *From Metaphysics to Ethics*, cit., pp. 44-45.

⁶ D.C. DENNETT, *Freedom Evolves*, cit., p. 225.

⁷ S. MORRIS, *The Impact of Neuroscience on the Free Will Debate*, in: «Florida Philosophical Review», vol. IX, n. 2, 2009, pp. 56-78, here p. 69.

⁸ D. BOURGET, D. CHALMERS, *What Do Philosophers Believe?*, in: «Philosophical Studies» vol. CLXX, n. 3, 2014, pp. 465-500.

⁹ See D. PEREBOOM, *Determinism Al Dente*, in: «Noûs», vol. XXIX, n. 1, 1995, pp. 21-45; D. PEREBOOM, *Living without Free Will*, cit.; D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit.

¹⁰ D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit., pp. 76-77.

¹¹ M. MCKENNA, *Conversation and Responsibility*, Oxford University Press, New York 2012; C. SARTORIO, *Causation and Free Will*, Oxford University Press, Oxford 2016.

¹² See R. CLARKE, *Dispositions, Abilities to Act, and Free Will: The New Dispositionalism*, in: «Mind», vol. CXVIII, n. 470, 2009, pp. 323-351.

¹³ See G. WATSON, *Two Faces of Responsibility*, in: «Philosophical Topics», vol. XXIV, n. 2, 1996, pp. 205-220; D. NELKIN, *Making Sense of Moral Responsibility*, Oxford University Press Oxford 2011; D. SHOEMAKER, *Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility*, in: «Ethics», vol. CXXI, n. 3, 2011, pp. 602-632.

¹⁴ See D.C. DENNETT, *Elbow Room*, MIT Press, Cambridge (MA) 1984; D.C. DENNETT, *Freedom Evolves*, cit.; J. LENMAN, *Compatibilism and Contractualism: The Possibility of Moral Responsibility*, in: «Ethics», vol. CXVII, n. 1, 2006, pp. 7-31; M. VARGAS, *Revisionism and Response to Fischer, Kane, and Pereboom*, cit.; M. VARGAS, *Building Better Beings*, Oxford University Press, New York 2013; B. VILHAUER, *Free Will Skepticism and Personhood as a Desert Base*, in: «Canadian Journal of Philosophy», vol. XXXIX, n. 3, 2009, pp. 489-511.

¹⁵ See D.C. DENNETT, *Elbow Room*, cit.; D.C. DENNETT, *Freedom Evolves*, cit.

¹⁶ See M. VARGAS, *Revisionism and Response to Fischer, Kane, and Pereboom*, cit.; M. VARGAS, *Building Better Beings*, cit.

¹⁷ See G. STRAWSON, *Freedom and Belief*, Oxford University Press, Oxford 1986; B. WALLER, *Freedom without Responsibility*, Temple University Press, Philadelphia 1990; D. PEREBOOM, *Determinism Al Dente*, cit.; D. PEREBOOM, *Living without Free Will*, cit.; D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit.

¹⁸ See I. KANT, *The Metaphysical Elements of Justice* (1797), translated by J. LADD, Bobbs-Merrill, New York 1963, pp. 99-107.

¹⁹ See D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit.; D. PEREBOOM, *Responsibility, Regret, and Protest*, in: D. SHOEMAKER (ed.), *Oxford Studies in Agency and Responsibility*, vol. 4, Oxford University Press, Oxford 2017, pp. 121-140.

²⁰ D. PEREBOOM, *Responsibility, Regret, and Protest*, cit.

²¹ P. HIERONYMI, *Articulating an Uncompromising Forgiveness*, in: «Philosophy and Phenomenological Research», vol. LXII, n. 3, 2001, pp. 529-554, here p. 546.

²² A. SMITH, *Moral Blame and Moral Protest*, in: N. TOGNAZZINI, D.J. COATES (eds.), *Blame: Its Nature and Norms*, Oxford University Press, New York 2013, pp. 27-48, here p. 43.

²³ See, e.g., I. HAJI, *Moral Appraisability*, Oxford University Press, New York 1998; M. HANSER, *Permissibility and Practical Inference*, in: «Ethics», vol. CXV, n. 3, 2005, pp. 443-470.

²⁴ For a related “conversational” account of blame see also M. MCKENNA, *Conversation and Responsibility*, Oxford University Press, New York 2012.

²⁵ See D. PEREBOOM, *Living without Free Will*, cit.; D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit.

²⁶ See D. PEREBOOM, *Living without Free Will*, cit., p. 177.

²⁷ See D. PEREBOOM, *A Defense of Free Will Skepticism: Replies to Victor Tadros, Saul Smilansky, Michael McKenna, and Alfred Mele on my Free Will, Agency, and Meaning in Life*, in: «Criminal Law and Philosophy», vol. XI, n. 3, 2017, pp. 617-636.

²⁸ See D. PEREBOOM, *Living without Free Will*, cit.; D. PEREBOOM, *Free Will, Agency, and Meaning in Life*, cit.