

Government House Compatibilism

Toby C.P. Solomon

Word count = 2800

September 17, 2018

Abstract

Many philosophers defend the idea that freedom of the kind required for moral responsibility is compatible with determinism. But there are also those who argue that another kind of freedom—the kind presupposed by rational decision-making—is not compatible with *belief* in determinism. The former—*responsibility-compatibilists*—had better hope that the latter—*deliberation-incompatibilists*—are wrong, because the conjunction of these two views leads us to an unwelcome position: *government house compatibilism*. This is the view that being both rational and morally responsible is compatible with determinism obtaining, but not with believing that it does, and that therefore we might have reason to obscure the truth about determinism. Debates about the relationship between determinism and moral responsibility are, therefore, more closely related to debates about the relationship between belief in determinism and rational decision-making than has, perhaps, been appreciated.

Discussions of free will often focus on its connection with moral responsibility.¹ In this vein compatibilists are those who believe:

Responsibility-Compatibilism: Freedom, of the kind required for moral responsibility, is compatible with determinism.

But free will is also connected to something else we care about: rational decision-making.² In this vein compatibilists are those who believe:

Deliberation-Compatibilism Freedom, of the kind presupposed by decision-making, is compatible with *belief* in determinism.³

Discussions of these two kinds of compatibilism have, at least in the last 60 years or so, mostly been conducted separately.⁴ However, there are important connections between them which we should not ignore, not least the one this paper is about: responsibility-compatibilism combined with deliberation-*in*compatibilism will lead us to some version of *government house compatibilism*—views on which we have reason to obscure the truth about determinism, and perhaps government house compatibilism itself, in order to preserve rational decision-making and/or moral responsibility. Responsibility-compatibilists should be concerned to avoid having to endorse this position.

¹Proof of this fact is hardly needed, but, if it is wanted, then it should suffice to note that McKenna and Pereboom's (2016) recent and comprehensive *Free Will: a Contemporary Introduction* dedicates 299 of its 304 pages to considerations primarily regarding moral responsibility.

²In the literature rational decision-making is usually referred to by the simpler, but much more ambiguous, term *deliberation* (hence the name deliberation-compatibilism). However, I think this use has led to a number of confusions that would be harder to make if we instead used the term *rational decision-making*. As such I will take this opportunity to begin reforming the terminology of the debate.

³For a (partisan) introduction to this debate see (Pereboom 2014, Ch. 5). Defenders of deliberation-compatibilism include: (Jones 1968), (Cowan 1969), (Dennett 1984), (Waller 1985), (Kapitan 1986, 1989), (Small 1988), (Pettit 1989), (Clarke 1992), (Bok 1998), (Nelkin 2004b, 2004a, 2011), (Levy 2006), (Pereboom 2008, 2014), (Nielsen 2011), and (Pendergraft 2014). Defenders of deliberation-incompatibilism (the negation of deliberation-compatibilism) include: (Ginet 1966), (Taylor 1964, 1966, 1968), (van Inwagen 1983), (Broadie 2001), (Henden 2010), (Stapleton 2010), and (Cohen 2017). Some who have discussed the issue but whose positions are more ambiguous include (Nozick 1969, 141), (Searle 2001), (Coffman and Warfield 2005), and (Fernandes 2016).

⁴Bok (1998) is a notable exception who explicitly connects her accounts of the two kinds of freedom. Several of the authors mentioned above discuss both topics in the same work, but their discussions tend to keep the issues rather separate.

Why be a Deliberation-Incompatibilist?

The motivation for responsibility-compatibilism need not be rehearsed here, but perhaps some motivation for deliberation-incompatibilism will be useful. Deliberation-incompatibilism is motivated by something like the following thought: decision-making is a necessarily comparative activity—I am always deciding whether to do *A or something else*—and so it is not possible to engage in rational decision-making unless I believe that there is more than one course of action actually open to me. But if determinism is true then there is only one course of action actually open to me. So, if I believe in determinism and its obvious consequences, then I must believe that there is only one course of action actually open to me and, therefore, cannot engage in rational decision-making.

This line of thought is also supported by one of most famous arguments in the debate about free will: the consequence argument. In van Inwagen's intuitive presentation:

If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born, and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (*including our present acts*) are not up to us (van Inwagen 1983, 16) [emphasis added].

Similarly Ginet's early presentation concludes that if determinism is true:

...no human being ever has a choice as to whether or not he shall behave as he does (Ginet 1966, 88).

These arguments may well tell us something about moral responsibility, but if anything surely they show that believing in determinism and its consequences is incompatible with rational decision-making—I cannot engage in rational decision-making about whether to do something I believe I do not have a choice about, or I believe is not up to me. Now that

we have some sense of what would motivate one to be a deliberation-incompatibilist we need to ask: what is so bad about government house compatibilism anyway?

The Problem with Government House Compatibilism

Endorsing both responsibility-compatibilism and deliberation-incompatibilism leads naturally to the following thought (and hence the name ‘government house compatibilism’⁵): even if I believe that determinism is true I should try to convince other people that *indeterminism* is true—because if they do come to believe in determinism they will no longer be able to engage in rational decision-making. And, in so far as I believe that rational decision-making generally leads to better results than other methods of settling what we will do, I seem to have strong reasons to want other people to engage in rational decision-making. This deception should, most likely, extend to myself in the future; I should now try to make it the case that my future self believes determinism is false.

These consequences might not be so bad if I was a responsibility-incompatibilist: what is the use, I might ask, of rational decision-making when no one is morally responsible for their acts anyway? My reason to obscure the truth about determinism is undercut by my belief that determinism would undermine moral responsibility anyway. And it seems that the responsibility-incompatibilist might have reason to obscure the truth about determinism that is not dependent on any link with rational decision-making.

But responsibility-compatibilists do generally believe that if determinism is true at least some people are sometimes morally responsible for their acts. Government house compatibilists should, therefore, do what they can to maintain our ability to engage in rational

⁵The analogy with a naïve version of Government House Utilitarianism should be clear. Government House Utilitarianism says that while Utilitarianism is in fact the true account of normative ethics, it itself says that the ethically best thing to do is believe some other theory. So those in the know about Utilitarianism should keep it to themselves and convince others to follow the other theory that Utilitarianism says it would be better to believe. There are no actual proponents of such a theory as far as I know, but the idea comes up in conversation often enough. Something along these lines is discussed by Sidgwick ([1907] 1962, 480-492) and Sen and Williams (1982, 15).

decision-making. Exactly what this requires them to do is a contingent matter, but it might, for example, require one to lie about the evidence one has regarding determinism, to destroy books which argue in favour of determinism, or to reduce funding to projects which aim to investigate determinism. These are not happy consequences.

The problem for government house compatibilists will be particularly bad if they subscribe to the common claim that moral responsibility is not only compatible with but in fact *requires* determinism.⁶ First, because this increases the tension within government house compatibilism. If determinism is necessary for moral responsibility, but belief in indeterminism is necessary for rational decision-making then one cannot be a responsible and rational decision-maker unless one believes a falsehood—that determinism is false when in fact it is true.

Second, because it gives this version of government house compatibilism reason to be *self-effacing*—that is, not only to obscure the truth of determinism but also to obscure the truth of government house compatibilism itself. If determinism is necessary for moral responsibility it will be impossible for any government house compatibilist to simultaneously engage in rational decision-making and believe that they are morally responsible. If they believe in determinism then they can believe they are morally responsible but cannot engage in rational decision-making; if they believe in indeterminism then they can engage in rational decision-making but cannot believe they are morally responsible.⁷ Such a government house compatibilist then has good reason to obscure the truth of determinism *and government house compatibilism itself* in so far as they have reason to regard belief in moral responsibility as important. Many compatibilists do think such belief is important, for example those who fear that a lack of belief in moral responsibility makes agents more likely to engage in immoral behaviour or undermine our legitimate punishing practices in

⁶See, for example, (Hume [1748] 2007, §VIII) or (Hobart 1934).

⁷Assuming that rational decision-making is compatible with indeterminism. Combining the view that rational decision-making is incompatible with both determinism and indeterminism with responsibility-compatibilism will obviously face even greater worries.

unacceptable ways (McKenna and Pereboom 2016, §11.8).

On the other hand those compatibilists who take moral responsibility to be compatible with both determinism and indeterminism (for example most Strawsonian compatibilists⁸) might try to take some of the sting out of the problem: if they have independent reason to deny that determinism is true they can claim that *in the actual world* there is no need for them to obscure the truth about anything. However, making this claim will undermine one of the reasons that either way compatibilisms of this kind are attractive: they make the existence of moral responsibility independent of how the world turns out to be with respect to determinism. It will also obscure what the difference is between this form of *compatibilism* and a libertarian-deliberation-incompatibilism: both insist that rational decision-making is incompatible with belief in determinism, and both rely on the falsity of determinism to make their account of moral responsibility plausible. Finally, if we are unsure whether determinism is true (as we surely in fact are) then these government house compatibilists might have good reason not to investigate any evidence that might lead us to believe in determinism, since their view will face the above challenges if we do come to believe (or even have high credence) in determinism. None of the above are consequences the responsibility-compatibilist should be happy to accept.

Making the Problem Worse

Things are not looking good for government house compatibilism, and they are about to get worse. Notice another very plausible connection between moral responsibility and rational decision-making:

Responsibility Requires Rational Decision-Making An agent, *X*, is morally responsible for ϕ ing only if *X*'s ϕ ing was a product of *X*'s rational decision-making, either the rational decision to ϕ or some prior rational decision leading to *X*'s ϕ ing.

⁸I.e. those following in the tradition of (Strawson 1962).

Intuitively this principle says that only those beings which can consider reasons for and against their acts and come to rational decisions about what to do are morally responsible.

Such a principle will be especially appealing to defenders of reasons-responsive compatibilism, for example Fischer and Ravizza (1998), who claim moral responsibility is dependent on acts being caused by mechanisms that are appropriately responsive to reasons—unimpeded rational decision-making is surely the example par excellence of such a mechanism.⁹ But it should also be attractive to other compatibilists: First, accounts of moral responsibility generally require that an agent was aware of both what they were doing and of its moral status (though these *epistemic conditions* are often mentioned only to be put to the side). But if an agent is incapable of rational decision-making it is unclear how such awareness is relevant to their acts, because it is unclear how such awareness could be involved in their production if not via rational decision-making. Second, acts which are not the products of rational decision-making are generally good examples of non-voluntary acts and compulsions. Anyone who wishes to deny the connection between moral responsibility and rational decision-making will have to give us some account of the difference between non-voluntary acts and acts which are not the product of any rational decision. Such an account is not likely to be easy to provide.

But when we combine the claim that moral responsibility requires rational decision-making with government house compatibilism we will find that an agent can be morally responsible for their acts only if they believe that determinism is false, whether or not it actually is. Again, this will be particularly bad if moral responsibility requires determinism. For then it will be impossible to be morally responsible while believing the truth; you are morally responsible only if determinism is true but you don't believe it. If undermining rational decision-making was not enough reason to obscure the truth of determinism, surely undermining rational decision-making *and* moral responsibility is.

⁹It should also be particularly attractive to Kantians who take morality to be grounded in practical reason (Bagnoli 2017, §2).

The Upshot

Responsibility-compatibilists should be interested in questions about determinism and rational decision-making more than they generally have been. They should be particularly interested to show that deliberation-compatibilism is true. This task will not necessarily be easy; the arguments and objections which feature in debates about responsibility-compatibilism are often quite different from those that feature in debates about deliberation-compatibilism. Not least because in the former discussion what is important is whether determinism obtains, while in the latter it is whether agents believe that determinism obtains. There are, of course, parallels between the debates, but the responsibility-compatibilist cannot be content with simply assuming that they will be able to show that deliberation-compatibilism is true—they need to actually show this, unless they are content to accept government house compatibilism.

To take a concrete example: Strawsonian compatibilism is based on the idea that moral responsibility is grounded in our reactive attitudes and that we will continue to feel the same reactive attitudes even if we come to believe that determinism is true. But nothing in this standard version of the view shows that we would not give up on rational decision-making were we to come to believe determinism were true. And if that is the case then the Strawsonian will, reactive attitudes notwithstanding, have to endorse government house compatibilism. Strawsonian compatibilists therefore need some argument for why deliberation-compatibilism is true, and it looks likely that this argument will be independent of the arguments they have advanced for their brand of responsibility-compatibilism. They have more work to do.

For another concrete example, consider Frankfurt's (1969) famous argument that alternative possibilities for action are not required for moral responsibility: the existence of a counterfactual intervener who blocks all but one course of action does not undermine our

intuition that an agent is morally responsible for following that course of action. This argument allows the responsibility-compatibilist to sidestep questions about whether more than one course of action is open in deterministic worlds. But it seems clear that if the agents in Frankfurt style cases *knew about the counterfactual intervener* then they would not be able to engage in rational decision-making about whether to follow the inevitable course of action the intervener wants them to—they could not believe that they had a choice about whether to do so. Though perhaps they can deliberate about related issues like whether they should follow this course of action for some reason or other, or whether they have reason to resist the intervener even if they will end up doing the same thing. The responsibility-compatibilist therefore cannot (at least without substantial modification) help themselves to Frankfurt’s arguments when discussing deliberation-incompatibilism—they need new tools.

The possibility of determinism poses a challenge for moral responsibility, while the possibility of belief in determinism poses a challenge for rational decision-making. While these issues are logically separable we cannot address them entirely separately: the plausibility of our response to one issue may depend on our response to the other. In particular, plausible responsibility-compatibilism relies on the truth of deliberation-compatibilism, and as such, responsibility-compatibilists should pay more attention to arguments concerning rational decision-making and (belief in) determinism.¹⁰

¹⁰Thanks to Al Hájek, Steven Manne, and Derk Pereboom for useful comments. This research was supported by an Australian Government Research Training Program Scholarship.

References

- Bagnoli, Carla. 2017. "Constructivism in Metaethics." In *The Stanford Encyclopedia of Philosophy*, Winter 2017, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.
- Bok, Hilary. 1998. *Freedom and Responsibility*. Princeton, N.J: Princeton University Press.
- Broadie, Sarah. 2001. "From Necessity to Fate: A Fallacy?" *The Journal of Ethics* 5:21–37.
- Clarke, Randolph. 1992. "Deliberation and Beliefs About One's Abilities." *Pacific Philosophical Quarterly* 73 (2): 101–113.
- Coffman, E. J., and Ted A. Warfield. 2005. "Deliberation and Metaphysical Freedom." *Midwest Studies in Philosophy* 29 (1): 25–44.
- Cohen, Yishai. 2017. "Deliberating in the Presence of Manipulation." *Canadian Journal of Philosophy*: 1–21.
- Cowan, J. L. 1969. "Deliberation and Determinism." *American Philosophical Quarterly* 6 (1): 53–61. JSTOR: 20009289.
- Dennett, Daniel C. 1984. *Elbow Room*. Cambridge, Massachusetts: MIT Press.
- Fernandes, Alison. 2016. "Varieties of Epistemic Freedom." *Australasian Journal of Philosophy* 94 (4): 736–751.
- Fischer, John Martin, and Mark Ravizza. 1998. *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge studies in philosophy and law. Cambridge ; New York: Cambridge University Press.
- Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66, no. 23 (December 4): 829.

- Ginet, Carl. 1966. "Might We Have No Choice." In *Freedom and Determinism*, edited by Keith Lehrer, 87–104. Studies in Philosophy. New York: Random House.
- Henden, Edmund. 2010. "Deliberation Incompatibilism." *Dialectica* 64 (3): 313–333.
- Hobart, R. E. 1934. "Free Will as Involving Determination and Inconceivable Without It." *Mind* 43 (169): 1–27. JSTOR: 2250169.
- Hume, David. (1748) 2007. *An Enquiry Concerning Human Understanding and Other Writings*. Edited by Stephen Buckle. Cambridge: Cambridge University Press.
- Jones, David. 1968. "Deliberation and Determinism." *Southern Journal of Philosophy*: 255–264.
- Kapitan, Tomis. 1986. "Deliberation and the Presumption of Open Alternatives." *The Philosophical Quarterly* 36 (143): 230–251.
- . 1989. "Doxastic Freedom: A Compatibilist Alternative." *American Philosophical Quarterly* 26 (1): 31–41.
- Levy, Neil. 2006. "Determinist Deliberations." *Dialectica* 60 (4): 453–459.
- McKenna, Michael, and Derk Pereboom. 2016. *Free Will: A Contemporary Introduction*. Routledge contemporary introductions to philosophy. New York, NY: Routledge.
- Nelkin, Dana Kay. 2004a. "Deliberative Alternatives." *Philosophical Topics* 32 (1): 215–40.
- . 2004b. "The Sense of Freedom." In *Freedom and Determinism*, edited by Joseph Keim Campbell, Michael O'Rourke, and David Shier, 105–134. Topics in Contemporary Philosophy. Cambridge, Massachusetts: MIT Press.
- . 2011. *Making Sense of Freedom and Responsibility*. Oxford: Oxford University Press.

- Nielsen, K. M. 2011. "Deliberation as Inquiry: Aristotle's Alternative to the Presumption of Open Alternatives." *Philosophical Review* 120, no. 3 (January 1): 383–421.
- Nozick, Robert. 1969. "Newcomb's Problem and Two Principles of Choice." In *Essays in Honor of Carl G. Hempel*, edited by Nicholas Rescher, 114–146. Dordrecht: Springer Netherlands.
- Pendergraft, Garrett. 2014. "Against Deliberation Restrictions." *Religious Studies* 50 (3): 341–357.
- Pereboom, Derk. 2008. "A Compatibilist Account of the Epistemic Conditions on Rational Deliberation." *The Journal of Ethics* 12, nos. 3-4 (September): 287–306.
- . 2014. *Free Will, Agency, and Meaning in Life*. Oxford : New York: Oxford University Press.
- Pettit, Philip. 1989. "Determinism with Deliberation." *Analysis* 49 (1): 42–44.
- Searle, John R. 2001. *Rationality in Action*. Cambridge, Massachusetts: MIT Press.
- Sen, Amartya, and Bernard Williams, eds. 1982. *Utilitarianism and Beyond*. Cambridge ; New York: Cambridge University Press.
- Sidgwick, Henry. (1907) 1962. *The Methods of Ethics*. 7th ed. London: Macmillan.
- Small, Robin. 1988. "Fatalism and Deliberation." *Canadian Journal of Philosophy* 18 (1): 13–30. JSTOR: 40231579.
- Stapleton, Sean. 2010. "Hard Incompatibilist Challenges to Morality and Autonomy." PhD diss., Cornell University.
- Strawson, Peter F. 1962. "Freedom and Resentment." *Proceedings of the British Academy* 48:1–25.

- Taylor, Richard. 1964. "Deliberation and Foreknowledge." *American Philosophical Quarterly* 1 (1): 73–80. JSTOR: 20009121.
- . 1966. *Action and Purpose*. Edgewood Cliffs, NJ: Prentice-Hall.
- . 1968. "Deliberation and Freedom." *Southern Journal of Philosophy* 6 (4): 265–268.
- Van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford : New York: Clarendon Press ; Oxford University Press.
- Waller, Bruce N. 1985. "Deliberating about the Inevitable." *Analysis* 45 (1): 48–52. JSTOR: 3327404.