

Countering Cova: Frankfurt-Style Cases are Still Broken

Neil Levy
Florey Institute of Neuroscience and Mental Health
University of Melbourne
Parkville, 3010
Australia
nllevy@unimelb.edu.au

Abstract: In his “Frankfurt-Style cases User Manual”, Florian Cova argues that my attempt to undermine Frankfurt-style cases fails. Cova advances two separate arguments. First, he distinguishes between agents’ possibilities and their capacities, and claims that my case establishes only that agents can gain or lose possibilities due to the presence of counterfactual interveners, but I need to establish that agents can gain or lose capacities for my argument to succeed. Second he distinguishes between direct and indirect versions of Frankfurt-style cases, and claims that my argument succeeds only against the latter. In response, I show that we have good reason to claim, contra Cova, that counterfactual interveners may alter agents’ capacities. I then show that my argument undermines direct variants of Frankfurt-style cases as effectively as indirect, by laying bare the nonconscious inferential processes that generate the intuition that the agents who feature in them are morally responsible.

Keywords: Moral responsibility; Alternative Possibilities; Capacities; Intuitions.

In his “Frankfurt-style cases user manual”, Florian Cova (2013) distinguishes two kinds of Frankfurt-style arguments against the principle of alternative possibilities (PAP), and argues that my attack on the soundness of Frankfurt-style cases succeeds, at most, only against one kind. Since either kind of argument can be used to undermine PAP, Cova suggests, the fact that my attack fails against at least one means that it does not succeed in rescuing PAP from the clutches of Frankfurt enthusiasts. In this brief response I will argue that my attack undermines both kinds of Frankfurt-style arguments. It shows that we are not entitled to rely on the intuitions that these arguments provoke, whether they are produced by the direct version of the Frankfurt-style argument or the indirect.

In a typical Frankfurt-style case (FSC), an agent performs a bad action for which, intuitively, they seem to be morally responsible, despite the fact that they lack alternative possibilities (see Frankfurt 1969 for the case which inspired all later variants). In an FSC, the action is one that another agent – canonically Black, a nefarious neuroscientist – wants the first agent to perform. Black has implanted a device in the agent’s brain that will fire iff the agent seriously consider performing some alternative action, causing them to perform the bad action. As it happens, though, the agent does not seriously consider performing some alternative action and the device does not fire. Because the agent performs the action on their own, using their entirely unimpaired capacities for reacting and responding to reasons, they seem to be morally responsible for what they do, but because the device would have fired and caused them to perform the bad action if they had seriously considered an alternative, they lacked alternative possibilities. Hence agents can be morally responsible despite lacking alternative possibilities, and PAP is falsified.

In the 2008 paper Cova cites, and elsewhere (Levy 2011; Levy 2012) I argued that we ought not to accept the soundness of the intuitions which (as I acknowledged) FSCs pump. I argued that these intuitions were produced by an internalist prejudice: an implicit belief that agents’ capacities can depend only on their own internal properties. I argued that *if* the agents who feature in FSCs are morally responsible, it is because they retain their responsibility-underwriting capacities. It is, I granted, intuitive that the agents who features in these cases retain these capacities: how could the *mere presence* of a counterfactual intervener, who by stipulation does not interfere in any way with what the agent does, affect the agent’s capacities? I then introduced a new kind of case, a Frankfurt-style enabling case (FEC) which, I claimed, produced an opposing intuition. In

an FEC an agent seems to *gain* a capacity due to the *mere presence* of a counterfactual intervener, who by stipulation does not interfere in any way with what the agent does. My prime FEC, modeled on Fischer and Ravizza's *Sharks* case (Fischer & Ravizza 1998), was *Phobia*:

Jillian is walking along the beach when she notices a child drowning. Jillian is a good swimmer, but she is pathologically afraid of deep water. She is so (internally) constituted that her phobia would prevent her from rescuing the child were she to attempt to; she would be overcome by feelings of panic. Nevertheless, she is capable of trying to rescue the child, and she knows that she is capable of trying. Indeed, though she knows that she has the phobia, she does not know just how powerful it is; she thinks (wrongly) that she could affect the rescue. Unbeknownst to Jillian, a good-hearted neurosurgeon has implanted her with a chip with which he monitors Jillian's neural states, and through which he can intervene if he desires to. Should Jillian decide (on her own) to rescue the child, the neurosurgeon will intervene to dampen her fear; she will not panic and will succeed, despite her anxiety, in swimming out to the child and rescuing her (Levy 2008, 234).

I argued that Jillian is (intuitively) responsible for failing to rescue the child (I note that no one who has discussed this case so far, including Cova, rejects this claim; apparently my intuition is widely shared). But she is morally responsible for her failure only if she has the capacity to rescue the child. And indeed, she has this capacity, though she has this capacity in part due to the presence of the counterfactual intervener. Thus what capacities we have may depend on features external to the agent, such as the mere presence of another agent who stands ready to help us to accomplish our goals. But, I claimed, if we can gain capacities in this kind of way, there is no principled reason to think that we cannot lose them in a parallel fashion. Since the intuition that agents are morally responsible in FSCs is vindicated only if the counterfactual intervener does not alter their capacities, the discovery that capacities can be altered by interveners who play no role in the actual sequence undermines the intuition; the intuition is shown to be without sufficient evidentiary value to play the role that the proponents of FSCs require of it.

Cova has two different lines of response to my argument. First, he argues that we ought to distinguish agents' *capacities* from their *possibilities* (2013, 4-5). He agrees that agents in FSCs lose possibilities due to the presence of the counterfactual intervener, but rightly notes that this should not be in dispute. FSCs are counterexamples to PAP *precisely because* the agents who feature in them lose possibilities. He also agrees that the agents in FECs may gain possibilities, but since no one disputes that merely

counterfactual interveners may make a difference to agents' possibilities, the fact that they play this role in FECs should not disturb enthusiasts for FSCs.

I agree with Cova that no one ought to be worried by the claim that merely counterfactual agents make a difference to agents' possibilities. I also agree with him that the fact that there is a possibility available for an agent does not entail that the agent has a correlative capacity. It is possible that I win the lottery but I do not have the capacity to win the lottery. But I deny that agents in FECs gain mere possibilities: rather they gain capacities. Indeed, I think that Cova requires this to be the case for his argument to succeed. Cova claims that the difference between Jillian in *Phobia* and her counterpart in a case that lacks only the presence of the counterfactual intervener, the difference that underwrites her moral responsibility, consists in the fact that she *omits* to save the child, something she can do only because she has the possibility of saving the child. But omitting to do something requires much more than the bare possibility of doing it. A (responsibility-grounding) omission is intentional and deliberate, and we can only omit to do things that we have the capacity to do (of course we can omit to *try* to do things we lack the capacity to do, but only if we have the capacity to try). I cannot *omit* to win the lottery, though I can omit to enter. Similarly, I cannot omit to sink a 40 foot putt, since I cannot intentionally sink such a putt though I can certainly omit to try.

Cova advances an argument that, if it succeeds, entails that my claim about the dependence of omissions on capacities fails. He presents us with the following variation on the *Jillian* thought experiment (2013, 5). In his case, the counterfactual intervener is absent but Jillian carries a potion that temporarily shuts down her phobia. Cova suggests that the possession of the potion gives Jillian the *possibility* of rescuing the child but – because she has to drink the potion before she rescues the child – not the capacity to rescue the child. Since Jillian seems just as responsible in Cova's variant of the thought experiment as in *Phobia* itself if she omits to save the child, mere possibilities are sufficient to ground moral responsibility, contrary to what I claimed above.

However, I see no good reason to accept the contention that Jillian has the possibility but not the capacity to rescue the child in Cova's modified case. In fact, Cova's case

begs the question against my claim that capacities can depend on features external to the agent. This is debate that is difficult to adjudicate, since there is no consensus on what having a capacity consists in, but certainly on many accounts Jillian-with-potion has the relevant capacity. Anyone who accepts a simple conditional analysis of capacities, according to which (roughly) an agent has the capacity to A just in case, were she to try to A she would probably succeed, should accept that the potion gives Jillian the capacity to rescue the child. Granted simple conditional analyses have come in for a lot of criticism, but successor theories, like dispositionalist accounts or restricted possibility accounts are compatible with the claim that the potion helps to ground her capacity (see Maier 2010 for discussion). Everything depends on how these accounts are developed. For a dispositionalist, for instance, it will depend on whether capacities depend on intrinsic properties of agents alone – a matter concerning which there is ongoing debate (see Clarke 2011 for discussion of how this debate bears on my argument against FSCs). It would be question-begging for Cova to simply assert that Jillian has a possibility but not a capacity in his modified case. Moreover, since Cova does not deny that Jillian gains a capacity in *Phobia* (2013, 6) – rather he denies that it is this fact that underwrites her responsibility – he is committed to taking my side on the debate whether extrinsic properties can ground capacities.

As I understand the dialectic of Cova's paper, he thinks that highlighting the capacity/possibility distinction is sufficient to block my argument against FSCs. But he has a second line of attack, which grants the soundness of my case if only for the sake of argument. According to this second line of attack, though my response may succeed against FSCs as Frankfurt (1969) envisaged them, they fail against another way of developing an FSC argument against PAP. Since Cova's second line of attack is conceived by him as being independent of the first, it needs a separate response.

Cova distinguishes two kinds of FSC arguments against PAP, *direct* and *indirect* (2013, 7-8). An indirect FSC depends for its soundness on a comparison between the agent who features in the actual case and an alternative case, which differs only in that the counterfactual intervener is absent. In an indirect FSC, the agent is held to be morally responsible *because* he does not differ in any relevant manner from a counterpart in a world that differs only in that the counterfactual intervener is absent. The indirect argument runs as follows (paraphrased from 2013, 3):

1. In an ordinary case, in which an agent performs a bad action by himself, the agent is morally responsible;
2. The only difference between the ordinary case and an FSC is the *mere* presence of a counterfactual intervener; but since the counterfactual intervener does nothing in the actual sequence, this difference cannot affect agents' moral responsibility.
3. Therefore the agent is morally responsible in an FSC, despite lacking alternative possibilities.
4. Therefore PAP is false.

Cova claims that if my argument succeeds, it is by undermining premise 2, the *no relevant difference* premise (2013, 3). That is, my argument shows (at most) that the mere presence of a counterfactual agent can make a difference to moral responsibility. But, Cova claims, the argument is impotent against the *direct* argument, because it dispenses with the *no relevant difference* premise. The direct argument runs as follows:

1. The agent is morally responsible in an FSC;
2. But due to the presence of the counterfactual intervener, the agent could not have done otherwise;
3. Therefore PAP is false.

I claim that my argument succeeds just as well against the direct version of an FSC as the indirect.

Cova cites experimental evidence that provides support for the claim that ordinary people understand FSCs in the way outlined by the direct argument, not the indirect (2013, 8-10). The claim strikes me as plausible. But I need not dispute it. After all, I explicitly acknowledged that FSCs produce the intuition that the agents who feature in them are morally responsible. 'Intuitions' are, roughly, immediate responses to cases: they are not the product of *conscious* inference. My argument was not that FSCs do not in fact produce the intuition that features in premise 1 of the direct argument: my claim was that the intuition is not *reliable*.

Intuitions have evidentiary value in philosophy, but no one thinks that the intuition that *p* is sufficient to establish that *p* is the case. The intuition that *p* is evidence for the truth of *p*, but the evidence is defeasible. I take myself to have offered a defeater of the intuition featuring in premise 1 of the direct argument. I argued, in effect, that the best

explanation of the intuition is that we engage in some kind of *nonconscious* inference. There is extensive evidence, both empirical and psychological, that we engage in nonconscious inferences after all. Consider the literature on cognitive dissonance. To make sense of much of the data, it seems necessary to attribute to the agent unconscious states and inferences over those states; for instance, to attribute to subjects an implicit belief with the content “I am a good person” and some kind of inference from that state (some kind of inference such as the following; *I am a good person; a good person doesn’t do things that are bad; what I just did isn’t all that bad*; see Aronson & Carlsmith 1962). There is also extensive evidence that unconscious inferences may produce *intuitions* without subjects having any access to the inferential processes at work. Cushman, Young & Hauser (2006) found that subjects’ moral intuitions were generated by commitment to implicit principles, some of which subjects could articulate and some of which they could not. Even such low-level processes as those involved in the generation of visual illusions such as the Müller-Lyer illusion seem to involve some kind of inferential processes.

In my 2008 paper and related work, I argued (following Fischer and Ravizza 1998) that the intuition cited in premise 1 of the direct argument is best explained by underlying commitments concerning the capacities of the agents who feature in FSCs. I then showed that these intuitions are unreliable insofar as they rely on what I called an internalist prejudice. This argument is as effective against the direct argument as the indirect: it explains away the intuition that the agent is morally responsible.

Cova and I agree on a great deal. We agree that counterfactual interveners can affect agents’ possibilities, and we agree that this is not sufficient to disturb defenders of PAP. Cova is also ready (albeit reluctantly) to concede that my case works against on variant of FSCs. But we disagree on crucial issues. We disagree concerning whether mere possibilities are sufficient to underwrite moral responsibility. More crucially, we apparently disagree about how intuitions are generated or about their evidential value. In this brief response, I have advanced considerations which I take to provide strong support for my side of these disputes.

Acknowledgements

Work leading to the production of this paper was generously supported by the Australian Research Council.

References

Aronson E, Carlsmith, JM (1962) Performance expectancy as a determinant of actual performance. *J Abnorm Soci Psychol* 65:178–182

Clarke R (2011) Responsibility, mechanisms, and capacities. *Mod Sch* 88:161-169

Cova F (2013) Frankfurt-style cases user manual: why Frankfurt-style enabling cases do not necessitate tech support. *Ethical Theory Moral Pract* doi: 10.1007/s10677-013-9456-x

Cushman FA, Young L, Hauser MD (2006) The role of reasoning and intuition in moral judgments: testing three principles of harm. *Psychol Sci* 17:1082-1089

Fischer JM, Ravizza M (1998) Responsibility and control: an essay on moral responsibility. Cambridge University Press, Cambridge

Frankfurt HG (1969) Alternate possibilities and moral responsibility. *J Philos* 66:829–839

Levy N (2008) Counterfactual intervention and agents' capacities. *J Philos* 105:223–239

Levy N (2011) Hard luck: how luck undermines free will and moral responsibility. Oxford University Press, Oxford

Levy N (2012) Capacities and counterfactuals: a reply to Haji and McKenna. *Dialectica* 66:607–620

Maier J (2010). Abilities. In: Stanford encyclopedia of philosophy. Available via <http://plato.stanford.edu/entries/abilities/>



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Levy, N

Title:

Countering Cova: Frankfurt-Style Cases are Still Broken

Date:

2014-01-01

Citation:

Levy, N. (2014). Countering Cova: Frankfurt-Style Cases are Still Broken. *Ethical Theory and Moral Practice*, 17 (3), pp.523-527. <https://doi.org/10.1007/s10677-013-9459-7>.

Persistent Link:

<http://hdl.handle.net/11343/283183>