



Article scientifique

Article

2014

Accepted version

Open Access

This is an author manuscript post-peer-reviewing (accepted version) of the original publication. The layout of the published version may differ .

---

## Frankfurt-Style Cases User Manual: Why Frankfurt-Style Enabling Cases Do Not Necessitate Tech Support

---

Cova, Florian

### How to cite

COVA, Florian. Frankfurt-Style Cases User Manual: Why Frankfurt-Style Enabling Cases Do Not Necessitate Tech Support. In: Ethical Theory and Moral Practice, 2014, vol. 17, n° 3, p. 505–521. doi: 10.1007/s10677-013-9456-x

This publication URL: <https://archive-ouverte.unige.ch//unige:109414>

Publication DOI: [10.1007/s10677-013-9456-x](https://doi.org/10.1007/s10677-013-9456-x)

## Dear Author

Here are the proofs of your article.

- You can submit your corrections **online**, via **e-mail** or by **fax**.
- For **online** submission please insert your corrections in the online correction form. Always indicate the line number to which the correction refers.
- You can also insert your corrections in the proof PDF and **email** the annotated PDF.
- For **fax** submission, please ensure that your corrections are clearly legible. Use a fine black pen and write the correction in the margin, not too close to the edge of the page.
- Remember to note the **journal title**, **article number**, and **your name** when sending your response via e-mail or fax.
- **Check** the metadata sheet to make sure that the header information, especially author names and the corresponding affiliations are correctly shown.
- **Check** the questions that may have arisen during copy editing and insert your answers/corrections.
- **Check** that the text is complete and that all figures, tables and their legends are included. Also check the accuracy of special characters, equations, and electronic supplementary material if applicable. If necessary refer to the *Edited manuscript*.
- The publication of inaccurate data such as dosages and units can have serious consequences. Please take particular care that all such details are correct.
- Please **do not** make changes that involve only matters of style. We have generally introduced forms that follow the journal's style.
- Substantial changes in content, e.g., new results, corrected values, title and authorship are not allowed without the approval of the responsible editor. In such a case, please contact the Editorial Office and return his/her consent together with the proof.
- If we do not receive your corrections **within 48 hours**, we will send you a reminder.
- Your article will be published **Online First** approximately one week after receipt of your corrected proofs. This is the **official first publication** citable with the DOI. **Further changes are, therefore, not possible.**
- The **printed version** will follow in a forthcoming issue.

### Please note

After online publication, subscribers (personal/institutional) to this journal will have access to the complete article via the DOI using the URL:

<http://dx.doi.org/10.1007/s10677-013-9456-x>

If you would like to know when your article has been published online, take advantage of our free alert service. For registration and further information, go to:

<http://www.springerlink.com>.

Due to the electronic nature of the procedure, the manuscript and the original figures will only be returned to you on special request. When you return your corrections, please inform us, if you would like to have these documents returned.

## Metadata of the article that will be visualized in OnlineFirst

1	Article Title	<b>Frankfurt-Style Cases User Manual: Why Frankfurt-Style Enabling Cases Do Not Necessitate Tech Support</b>	
2	Article Sub- Title		
3	Article Copyright - Year	<b>Springer Science+Business Media Dordrecht 2013 (This will be the copyright line in the final PDF)</b>	
4	Journal Name	Ethical Theory and Moral Practice	
5		Family Name	<b>Cova</b>
6		Particle	
7		Given Name	<b>Florian</b>
8	Corresponding Author	Suffix	
9		Organization	University of Geneva
10		Division	Swiss Centre for Affective Sciences
11		Address	7 rue des Battoirs, Geneva 1205, Switzerland
12		e-mail	florian.cova@gmail.com
13		Received	
14	Schedule	Revised	
15		Accepted	9 August 2013
16	Abstract	<p>'Frankfurt-style cases' (FSCs) are widely considered as having refuted the Principle of Alternate Possibilities (PAP) by presenting cases in which an agent is morally responsible even if he could not have done otherwise. However, Neil Levy (J Philos 105:223–239, 2008) has recently argued that FSCs fail because we are not entitled to suppose that the agent is morally responsible, given that the mere presence of a counterfactual intervener is enough to make an agent lose responsibility-grounding abilities. Here, I distinguish two kinds of Frankfurt counter-arguments against the PAP: the direct and the indirect counter-arguments. I then argue that Levy's argument, if valid, can shed doubt on the indirect argument but leaves the direct argument untouched. I conclude that FSCs can still do their job, even if we grant that the mere presence of a counterfactual intervener can modify an agent's abilities.</p>	
17	Keywords separated by ' - '	Moral responsibility - Frankfurt cases - Abilities - Alternative possibilities - Experimental philosophy	
18	Foot note information		

**Frankfurt-Style Cases User Manual: Why Frankfurt-Style  
Enabling Cases Do Not Necessitate Tech Support**

**Florian Cova**

Accepted: 9 August 2013  
© Springer Science+Business Media Dordrecht 2013

**Abstract** ‘Frankfurt-style cases’ (FSCs) are widely considered as having refuted the Principle of Alternate Possibilities (PAP) by presenting cases in which an agent is morally responsible even if he could not have done otherwise. However, Neil Levy (J Philos 105:223–239, 2008) has recently argued that FSCs fail because we are not entitled to suppose that the agent is morally responsible, given that the mere presence of a counterfactual intervener is enough to make an agent lose responsibility-grounding abilities. Here, I distinguish two kinds of Frankfurt counter-arguments against the PAP: the direct and the indirect counter-arguments. I then argue that Levy’s argument, if valid, can shed doubt on the indirect argument but leaves the direct argument untouched. I conclude that FSCs can still do their job, even if we grant that the mere presence of a counterfactual intervener can modify an agent’s abilities.

**Keywords** Moral responsibility · Frankfurt cases · Abilities · Alternative possibilities · Experimental philosophy

For a long time, the Principle of Alternate Possibilities (PAP) has dominated the philosophical debate about the compatibility of moral responsibility with determinism. The PAP states that:

(PAP) Persons are morally responsible for what they have done only if they could have done otherwise.

Under a certain interpretation, one can deduce from PAP the incompatibility of moral responsibility with determinism for, under a certain interpretation of “could have done otherwise”, an agent in a deterministic world cannot do otherwise. Of course, compatibilist philosophers (who argue for the compatibility of moral responsibility with determinism) have denied that this interpretation of the ability to do otherwise is the one relevant for moral responsibility, while incompatibilists (who argue for the incompatibility of moral responsibility with determinism) have argued it is.

It is widely held that this “dialectical stalemate” (Fischer 1994) was broken in 1969 with the publication of Harry Frankfurt’s seminal paper ‘Alternate Possibilities and Moral Responsibility’ (Frankfurt 1969). In this paper, Frankfurt proposed a recipe to create what came to be called ‘Frankfurt-style cases’ (FSC), that is: cases in which an agent is morally

---

F. Cova (✉)  
Swiss Centre for Affective Sciences, University of Geneva, 7 rue des Batoirs, 1205 Geneva, Switzerland  
e-mail: florian.cova@gmail.com

responsible *even if* he could not have done otherwise. This is achieved by constructing cases in which an agent acts entirely *on his own*, but in which a *counterfactual intervener* would have forced him to act the same way if he hadn't. Thus, FSCs constitute counter-examples to the PAP and suggest that the ability to do otherwise might not be among the necessary conditions for moral responsibility.

However, not everyone has been convinced by FSCs' ability to defuse the PAP. Some have argued that agents in FSCs still have alternate possibilities.<sup>1</sup> Others have argued that FSCs presuppose determinism and hence that claiming that agents in FSCs are morally responsible only begs the question against incompatibilism.<sup>2</sup> More recently, however, Neil Levy (2008) has argued that FSCs fail because they suppose that the presence of a counterfactual intervener has no impact on the agent's responsibility-grounding capacities. But Levy provides examples suggesting that the mere presence of a counterfactual intervener can be enough to make an agent gain responsibility-grounding capacities. If so, then how could we be entitled to presuppose that it cannot also make an agent lose such capacities?

A possible line of answer is to deny that the mere presence of a counterfactual intervener is enough for an agent to lose responsibility-grounding capacities (Haji and McKenna 2011). However, this is not the path I will take in this paper. Here, I will argue that, even if one grants Levy that the mere presence of a counterfactual intervener is enough for an agent to lose responsibility-grounding capacities, this won't save the PAP from FSCs, for there are two distinct ways FSCs can be used against the PAP, and Levy's argument only works against one of them.

### 1 Levy's 'Frankfurt-Style Enabling Cases'

Let's first focus on Levy's argument. Here is a standard FSC<sup>3</sup>:

*Voting*: When she enters the voting booth, Connie has not yet made up her mind whether to vote Democrat or Republican. Unbeknownst to her, an evil but gifted neuroscientist monitors her neural states using a computer chip he has implanted in her brain. The chip gives the neuroscientist the power to intervene in Connie's neural processes. Connie is psychologically constituted so that it is a necessary (though not sufficient) condition of her voting Democrat that she thinks deeply about a certain policy; were Connie to think deeply about this policy, the neuroscientist would intervene, causing her to choose to vote Republican. But Connie does not think deeply about the policy and proceeds to vote Republican on her own. The neuroscientist and his device play no role in bringing about Connie's vote.

This case is supposed to show that Connie can freely vote Republican (and be morally responsible for her vote) even if she couldn't have done otherwise (because of the neuroscientist's chip). The argument is supposed to go like this: imagine the same case, except that there is no evil neuroscientist. In this case (let's call it *Normal Voting*), Connie acts normally, so that she's free and morally responsible according to both compatibilist and incompatibilist criteria. Now, let's compare *Normal Voting* and *Voting*: in *Voting*, Connie acts exactly the same way as in *Normal Voting*, since the evil neuroscientist does not act. Surely, the mere

<sup>1</sup> This is known as the *flickers of freedom* defense. See for example: van Inwagen (1983).

<sup>2</sup> This is known as the *dilemma defense*. See for example: Ginet (1996).

<sup>3</sup> I use Levy's own example. It is a variation on a case created by Fischer (2006) and improved by Pereboom (2001).

presence of the neuroscientist cannot have an impact on the way Connie acts and the way she came to act this way. But, if the mere presence of the neuroscientist does not change anything as far as Connie is concerned, we must conclude that Connie is responsible in *Voting*, for she was so in *Normal Voting*. However, Connie could not have done otherwise. Therefore, Connie is morally responsible *even if* she could not have done otherwise, and this counts as a counter-example to the PAP. 82-87

More generally, the basic ‘Frankfurt counter-argument’ against the PAP can be summarized in the following way. First you have a *normal* case in which an agent is morally responsible for his action, then you have a *Frankfurt* case that is similar to the normal case except that a counterfactual intervener lurks in the shadows and would have forced the agent to act the way he did if he had not. Then, the argument runs this way: 88-92

1. The agent is morally responsible in the *normal* case. 93
2. Since the only difference between the *normal* case and the *Frankfurt* case is the presence of a counterfactual intervener, and since this intervener does not intervene in the Frankfurt case, there is no difference between the *normal* and *Frankfurt* cases as far as the agent’s moral responsibility is concerned. 94-97
3. Thus, the agent is morally responsible in the *Frankfurt* case. 98
4. However, due to the presence of the counterfactual intervener, the agent could not have done otherwise in the *Frankfurt* case. 99-100
5. Then, the PAP is false, since the agent in the *Frankfurt* case is morally responsible even though he could not have done otherwise. 101-102

Levy rejects this counter-argument by arguing against (or, more precisely, casting doubt upon) premise 2. More precisely, Levy argues that it is not clear that the mere presence of a counterfactual intervener makes no difference, for it might be that his mere presence is enough to make the agent lose some responsibility-grounding abilities. Thus, the inference from the fact that the agent is responsible in the *normal* case to the fact that he is responsible in the *Frankfurt* case is not warranted. 103-108

However, Levy does not show directly that the mere presence of a counterfactual intervener can cause the loss of responsibility-grounding abilities. His strategy consists in introducing new cases, the ‘Frankfurt-style enabling cases’ (FECs), in which the mere presence of a counterfactual intervener is enough to make an agent *gain* responsibility-grounding abilities, and then arguing that if the mere presence of counterfactual intervener is enough to make an agent *gain* responsibility-grounding abilities, then there is no reason to think that it cannot also make an agent *lose* responsibility-grounding abilities. 109-115

Here is an example of FEC: 116

*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 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. 118-129

Imagine that Jillian does not try to save the child and the child drowns: she would be responsible for the child's death, for she could have saved him if she tried and thought she could. But what if she had let the child drown in a similar scenario *without* the presence of a good-hearted neurosurgeon: surely, in this case, she would not have been responsible for the child's death.<sup>4</sup> According to Levy, this means that the mere presence of the neurosurgeon is enough to make Jillian morally responsible for the child's death, that is, to make her *gain* responsibility-grounding abilities. It then seems possible that the mere presence of a counterfactual intervener in FSCs can be enough to make the agent *lose* responsibility-grounding abilities.

There seems to be two main ways of rejecting Levy's argument. The first is to deny that agents in FECs actually *gain* responsibility-grounding abilities. The second is to grant it, but to deny that it implies that the mere presence of a counterfactual intervener can also make an agent *lose* responsibility-grounding abilities in FSCs. In their recent responses to Levy, Haji and McKenna (2011) adopt the former solution, while Clarke (2011) defends the latter. Personally, had I to choose, I would go for a combination of both ways: either agents do not gain abilities in FECs, or their losing such abilities in FSCs is not sufficient for impacting their moral responsibility because it requires an additional factor that is not to be found in FSCs. Here is how the argument would go:

1. Let's begin by examining Levy's method: how does he conclude from cases such as *Phobia* that the mere presence of a counterfactual intervener makes the agent gain some ability? The argument is the following: because Jillian is responsible for the child's death in *Phobia* but not in an equivalent case without counterfactual intervener (let's call this case *Normal Phobia*), there must be some factor explaining this difference. According to Levy, the best explanation is that Jillian in *Phobia* has gained a responsibility-grounding ability.

However, it is not the only possible explanation: the difference in Jillian's moral responsibility for the child's death between *Phobia* and its normal counterpart could also be explained by the fact that Jillian in fact omits to do something different in the two cases. Indeed, it should be noted that an agent's responsibility (and blameworthiness) does not depend only on his responsibility-grounding abilities. For example, an agent's responsibility also depends on what he actually *does*. Thus, two agents with equal responsibility-grounding abilities might end up with different levels of responsibility in a person's death if the first tried to save this person and failed while the second stabbed this same person repeatedly in the back. This may seem (and is) trivial, but this means that one cannot directly infer a difference in responsibility-grounding abilities from a mere difference in responsibility, for this difference could be due to a difference in something else, for example in what the agent did.

Now, one might argue that Jillian actually does the same thing in *Phobia* and in *Normal Phobia* (that is: nothing, or at least deciding to do nothing) and thus that the only remaining option is that the difference in moral responsibility between the two cases is due to changes in Jillian's responsibility-grounding abilities. But that would be jumping to the conclusion, for there is something else that differs between the two cases: namely, what Jillian *omits to do*. Indeed, a difference in moral responsibility can also be

<sup>4</sup> Granted, she would still deserve blame for not trying to save the child while she thought she could, but this blame would be directed at what her omission reveals about her, and not a result of her being responsible for the child's death.

explained by a difference in what someone omits to do (in the broad sense of what someone *could have done* and *fails* to do, and not in the narrow sense of what someone *deliberately chooses* not to do): two individual with equal responsibility-grounding might end up with different levels of responsibility in a person's death if the first omitted to save this person while the second did not (because, for example, one had the possibility to save the person while the other was far away and not aware that the person was in danger).

Thus, one might say that the difference in Jillian's moral responsibility between *Phobia* and *Normal Phobia* is not due to a difference in Jillian's responsibility-grounding abilities, but to a difference in what Jillian omits to do. Indeed, in *Normal Phobia*, Jillian's decision to let the child drown leads to an omission *to try to help*, but not to an omission *to help* for, had she tried to help the child, she could not have succeeded. On the contrary, in *Phobia*, Jillian's omission to try to help also leads to an omission to help, for, had she tried to save the child, she would have succeeded. In other words, if Jillian's moral responsibility differs between both cases, it is because what Jillian could have done and fails to do (and thus, the nature of her omission) is also different. As a result, we do not need Jillian gaining abilities to account for this difference, and Levy's conclusion from Jillian gaining moral responsibility to her gaining responsibility-grounding abilities does not follow.

However, one could ask why the nature of Jillian's omission changes when there is a counterfactual intervener: is it not precisely because the presence of a counterfactual intervener makes Jillian gains a new ability? Not necessarily, as one can say that the mere presence of a counterfactual intervener changes the nature of her omission by opening to her a new *possibility* (the one of saving the child). More precisely, one could say that the fact that Jillian would have gained a new *ability* in case she had tried to save the child is enough to open to her a new *possibility* in the actual case, which is enough for turning her mere omission to try to help into an omission to help. And it is this change in Jillian's omission that drives the change in responsibility ascriptions, for one cannot be responsible for the death of a person one did not kill if one did not omit to prevent her death.

This distinction between *ability* and *possibility* can be illustrated using the following case: imagine there is no good-hearted neurosurgeon but that Jillian, being aware of her phobia, carries with her a potion that temporarily shuts down her phobia. It seems intuitive to say that carrying this potion with her does not grant her the *ability* to save the child (since she would have to drink the potion to gain this ability) but the *possibility* to save the child (since there is a possible world open to her in which she can gain this ability). In such a case, we have an agent gaining a possibility without gaining some ability. Inversely, *mask* cases are cases in which an agent (or an object) retains an ability while possibilities disappear (thus, in the now well-known philosophical example of the fragile cup that has been enchanted never to break, the cup keep its ability to break—fragility—but has lost all possibilities of doing so).

Thus, one might argue that we cannot infer from agents gaining moral responsibility in FECs that they also gained responsibility-grounding abilities, for there is another explanation according to which what they gain are in fact new possibilities. If this hypothesis is right, then we can explain why the mere presence of a counterfactual intervener can have an impact on the agent's moral responsibility in FECs but not in standard FSCs. FSCs traditionally involve commissions (like voting Republican) whose nature depends only on the actual course of events, and not on what could happen in counterfactual scenarios (and on what other possibilities are opened to the agent). Thus, the mere presence of a counterfactual

intervener does not matter in FSCs, since its presence does not affect the nature of the agent's action.<sup>5</sup>

2. So far my argument has been that one cannot deduce from the fact that agents in FECs gain moral responsibility that they gain new responsibility-grounding abilities, for it might be explained simply by the fact that they gain new relevant possibilities. However, one might argue that, in Levy's FECs, it is very clear that agents gain new abilities, and that we do not have to deduce it from their gaining moral responsibility. Take *Phobia*: isn't it clear that, in this case, Jillian gains the ability to save the child, or to resist her phobia? How one could deny that Jillian *did* gain these abilities by the mere presence of the counterfactual intervener?

But I do not want to deny that Jillian gained such abilities. I do not even need to. I just want to ask: are these abilities "responsibility-grounding abilities"? In a very strong sense in which "responsibility-grounding abilities" are abilities that are *always* necessary for moral responsibility (let's call these *essential* responsibility-grounding abilities), the answer is: no. Clearly, one can be morally responsible for one's actions (for example, stealing a book in a public library) even if one does not have the ability to save a drowning child or to resist one's phobia of deep water. Thus, one might wonder why we should be worried that agents in FSCs lose such abilities. Now, one might say that these abilities are "responsibility-grounding abilities" in the sense that they are required for Jillian to be morally responsible for a *particular* action or omission in *particular* circumstances (let's call them *circumstantial* responsibility-grounding abilities). But, given the imagination of philosophers, any capacity can be considered to be such a circumstantial responsibility-grounding ability. More importantly, as we have seen, losing such circumstantial responsibility-grounding abilities is not enough to be deprived of one's moral responsibility; it only happens in the relevant circumstances, and though Jillian is not morally responsible for the child's death in *Normal Phobia*, she would surely be responsible for stabbing a woman walking on the beach.

This leads to the following question: when is the gain or loss of such circumstantial responsibility-grounding abilities relevant to moral responsibility? Once again, the answer is: in circumstances where this gain or loss changes the range of possibilities offered to the agents. The ability to save the child (or to resist phobias) is relevant to Jillian's moral responsibility only because it gives her the possibility to save the child. In circumstances in which these abilities do not change the agent's possibilities, they are irrelevant to assessing

<sup>5</sup> This explanation of Levy's FECs directly leads to the following predictions: that FECs can only be cases in which the very nature of what the agent does or omits to do depends on what the agent can achieve in counterfactual scenario. Omissions are a paradigmatic case, since they essentially depend on what the agent could have done. Thus, one would expect to find FECs more easily in cases involving omissions, and this would explain why Levy has indeed proposed only cases involving omissions.

Such considerations might lead one to the stronger prediction that FECs will be only found in cases of omissions and never in cases of commissions. However, this more particular prediction rests on the plausible but controversial claim that the nature of commissions never depends on what the agent could have done in counterfactual scenarios—a claim I am not certain of. Imagine that Peter tries to kill Jones but Jones shoots Peter before Peter can fulfill his murderous intention. Imagine also the Peter was also persuaded that it was the only way to save his own life, but that lurking in the shadow was Black, an evil neuroscientist who would have taken control over Peter and prevented him from killing Jones if Jones had only uttered "Frankfurt". If one admits that killing in self-defense means killing while there was no softer way of saving oneself, one could say that the presence of this counterfactual intervener is enough to transform Jones' 'killing in self-defense' into a simple 'killing'. However, this is no problem for my defense of FSCs, for all I need is that there are FSCs in which the nature of the agent's action does not depend on what other possibilities are opened to him.

For an objection to Levy that rests precisely on the distinction between commission and omission, and concludes that the lesson FECs draw from omissions cannot be transposed to actions in FSCs, see Clarke (2011).

this agent's moral responsibility. But, once again, this means that the possible loss of such circumstantial responsibility-grounding in FSCs is no threat to the agent's moral responsibility: agents in FSCs have already only one possibility open to them. That they lose this only possibility is excluded by the very formulation of FSCs. So, in FSCs, there is no possibility to be lost, and the potential loss of circumstantial responsibility-grounding abilities is irrelevant to moral responsibility.<sup>6</sup>

3. Finally, one might argue that the agent in FECs does not only gain circumstantial responsibility-grounding abilities, but also essential responsibility-grounding abilities. If this were really the case, then the possibility that agents in FSCs lose essential responsibility-grounding abilities would indeed be threatening. But one may ask: which essential responsibility-grounding abilities are gained in FECs? What essential responsibility-grounding abilities does Jillian gain in *Phobia*? The most plausible answer is that Jillian gains the ability to respond to reasons, rather than being overcome by phobia. However, it seems wrong to say that Jillian gains such an ability. Surely, in everyday life, Jillian does respond to reasons. Her phobia only temporarily *masks* this ability in particular circumstances. Thus, what the presence of the counterfactual intervener does is not offering Jillian a new essential responsibility-grounding ability, but preventing one she already has from being masked. At best, the mere presence of the counterfactual intervener preserves one of Jillian essential responsibility-grounding abilities, but does not make new ones appear.

Thus, it seems that we have reasons to doubt Levy's argument against FSCs: either agents in FECs do not gain responsibility-grounding abilities, or the possible loss of these abilities does not threaten the responsibility of agents in FSCs because their doing so depends on the agent losing possibilities, something that cannot happen to agents in FSCs. However, I will not press this point further, for the main point of this paper will be that, even if we grant Levy his point and accept that we are not entitled to say that the mere presence of a counterfactual intervener does not change anything in FSCs, FSCs can still be used against PAP, for there are at least two different ways in which FSCs can be used to undermine PAP.

## 2 Two Ways of Using Frankfurt-Style Counterexamples 284

As I have introduced it in the previous section, the basic Frankfurt counter-argument works in an *indirect* way. This is why I will call this version of the counter-argument the *indirect* Frankfurt counter-argument (henceforth: IFA). IFA starts with an intuition about a certain case (the *normal* case) and then proceeds by transferring this intuition to a second case (the *Frankfurt* case) on the assumption that there is no significant difference between the two cases along the relevant dimensions. Thus, IFA can be compared to the Manipulation Argument that starts with our intuition that a manipulated agent is not morally responsible

<sup>6</sup> One might be tempted to say that Levy just has to translate his argument from the language of "abilities" to the language of "possibilities" and argue that, because the mere presence of a counterfactual intervener is enough for the agent to gain possibilities if FECs, then it is possible that the mere presence of a counterfactual intervener is enough to make the agent lose possibilities in FSCs. However, this is no objection to FSCs because it is exactly what FSCs presuppose: that the mere presence of the counterfactual intervener makes the agent lose the possibility to do otherwise. This objection would only work if such an argument showed that agents in FSCs lose the very possibility of doing what they have to do, which would be a very strange conclusion.

for his actions and tries to transfer this intuition to the case of an agent in a deterministic world, on the ground that there is no relevant difference between the two cases. 292  
293

Granted, there are more steps in the standard Manipulation Argument.<sup>7</sup> However, there is another important difference between the standard Manipulation Argument and IFA: while we may not have the direct intuition that an agent in a deterministic universe is not morally responsible for his action,<sup>8</sup> we *do* have the direct intuition that an agent in a FSC is morally responsible for his action even if he could not have done otherwise. What I mean here by ‘direct intuition’ is that this intuition is not dependent on the comparison of the FSC with another case, but directly triggered by reading and considering the FSC itself. For example, I contend that, in *Voting*, we immediately have the intuition that Connie is responsible for her Republican vote, even without making the comparison with the *Normal Voting* case. 294  
295  
296  
297  
298  
299  
300  
301  
302

For those who would (legitimately) doubt such a claim about ‘our’ intuitions, it is possible to ground this claim on empirical data. Miller and Feltz (2011) have given the following vignette to a certain number of participants: 303  
304  
305

*Car:* Mr. Green wants Mr. Jones, the security guard, to steal Mrs. Green’s car at 12:00 am on October 7. However, Mr. Green doesn’t entirely trust Mr. Jones to do the job, so he has taken some extraordinary measures. Mr. Green has consulted neuroscientists who have implanted a device in Mr. Jones’s brain without Mr. Jones’s knowledge. This device has isolated the “decision-making” neurons in Mr. Jones’s brain and is programmed to send, at exactly 12:00 am, impulses that will certainly cause Mr. Jones to decide to steal the car just then. However, as it happens, at exactly 12:00 am, Mr. Jones decides on his own to steal the car and does it. Since Mr. Jones decides on his own to steal the car, the impulses from the device were ineffectual because the decision-making neurons were activated by the decision-making process of Mr. Jones himself. However, if Mr. Jones had not, just then, decided on his own to steal the car, the device would have activated his decision-making neurons, and Mr. Jones would have decided to steal the car anyway. 306  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320

Overall, participants tended to consider Mr. Jones morally responsible for having decided to steal the car, and whether they considered Mr. Jones able to do otherwise had no influence on their judgment. Given that the vignette is presented alone, it seems that most people do have a *direct* intuition that agents are morally responsible in FSCs. 321  
322  
323  
324

However, as a reviewer for this journal rightly pointed out, this is not enough evidence. Indeed, it could be possible that most participants came to this judgment by mentally comparing this vignette with a counterfactual scenario in which Mr. Green is absent (thus reproducing IFA at a purely internal level). To test this hypothesis, I reproduced Miller and Feltz’s experiment by giving the same vignette to 200 participants<sup>9</sup> and asking them the following questions: 325  
326  
327  
328  
329  
330

1. Was it possible for Mr. Jones to avoid stealing the car at 12:00 am on October 7? (YES/NO) 331  
332
2. At 12:00 am on October 7, could Mr. Jones have done anything other than steal the car just then? (YES/NO) 333  
334
3. Is Mr. Jones morally responsible for stealing the car? (YES/NO) 335

<sup>7</sup> Three, for example, in Pereboom’s version. See Pereboom (1995).

<sup>8</sup> More on this in Section 3.

<sup>9</sup> Participants were recruited through Amazon Mechanical Turk and paid \$0.3 for their participation. All lived in United States. Mean age was 31.93 and 79 were women.

- 4. About your answer to Question 3: describe how you reached this conclusion and what made you think so. 336  
337
- 5. To answer Question 3, did you mentally compare this case to a case in which Mr. Jones would have acted the same way but in which Mr. Greene would not have been present? (YES/NO) 338  
339  
340
- 6. Did you have to think a lot to answer Question 3, or was it easy? (I had to think a lot/It was easy) 341  
342

Questions 1 and 2 were control questions: I was only interested in participants who considered the case as a case in which Mr. Jones could not have done otherwise, and thus gave a 'NO' answer to both questions. 179 participants out of 200 did so, and only their answers are analyzed here. 343  
344  
345  
346

Question 3 tested whether these participants considered Mr. Jones morally responsible for stealing the car though he could not have done otherwise. Out of the 179 remaining participants, 163 answered 'YES'. Thus, most participants (91 %) expressed the intuition that an agent could be morally responsible for his action even though he could not have done otherwise. 347  
348  
349  
350  
351

Now, is this intuition directly elicited by considering the case, or does it require its comparison with a 'normal' case without counterfactual intervener? Questions 4 and 5 are designed to answer this question. Let's begin with Question 5, which directly asks participants whether they did such a mental comparison. Out of 179 these participants, 82 (46 %) answered 'YES'. This means that more than half of participants seem not to have made such a comparison. If such a comparison were necessary for eliciting the intuition that Mr. Jones is morally responsible for stealing the car even though he could not have done otherwise, then one would expect participants that did not make such a comparison to attribute less moral responsibility to Mr. Jones. However, though 11 % of participants who did *not* do the comparison denied moral responsibility while only 6 % of those who *did* the comparison did so, this difference turned out not to be significant.<sup>10</sup> Thus, we have no evidence that whether a participant actually made the comparison led him to attribute more moral responsibility to the agent.<sup>11</sup> Moreover, it also seems that, in absence of such a comparison, participants still tend to attribute moral responsibility to the agent, since 89 % of participants who declared *not* making the comparison still considered the agent to be morally responsible (Table 1). 352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366Q2

Still, one could say that, for 46 % of the participants, the intuition that Mr. Jones is morally responsible comes from such a comparison. However, this claim is not warranted by my results; although 46 % of participants did claim to have made this comparison, they did not say that this was what led them to judge Mr. Jones responsible. Question 4 was precisely designed to investigate what led participants to judge Mr. Jones responsible. Focusing on the 163 participants who thought Mr. Jones to be responsible, I submitted their answers to two independent coders (EP and VS). For each answer, coders had to determine whether the explanation given by the participant featured one or more of the following categories of explanations: (i) **Choice**: participants stress the fact that Mr. Jones made a *decision* or a *choice*, (ii) **On his own**: participants stress the fact that Mr. Jones decided or acted *on his own*, (iii) **Not forced**: participants stress the fact Mr. Jones was *not caused to act by external forces*, (iv) **Comparison 1**: participants compare the current case to a case *in which Mr. Green is not even present*, and stress the fact that Mr. Jones *would have acted the same* 367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379

<sup>10</sup>  $\chi^2(1, N=179)=0.93, p=.34$ .

<sup>11</sup> Further reasons to doubt this claim come from a pilot study for that study in which we found that participants who declared having done the comparison were (also non-significantly) more likely to deny moral responsibility to the agent.

t1.1 **Table 1** Participants' answers for Question 3 (*Responsibility*) and Question 5 (*Comparison*)

t1.2	<i>Number of participants who:</i>	Judged that Mr. Jones was morally responsible	Judged that Mr. Jones was <i>not</i> morally responsible
t1.3	Did the comparison with a 'normal' case	77 (43 %)	5 (3 %)
t1.4	Did <i>not</i> do the comparison with a 'normal' case	86 (48 %)	11 (6 %)

way in this case, and (v) **Comparison 2:** participants compare the current case to a case in which Mr. Green activates the device, and stress the fact that the current case is different from this one.<sup>12</sup>

Inter-raters agreement was fairly substantial.<sup>13</sup> Table 2 indicates for both coders and the trials on which they agree the number of explanations that fitted each category. As can be seen, very few participants gave as a reason that the agent would have acted in the very same way had Mr. Green not been there, or that the presence of Mr. Green's device made no difference (category iv). Rather, the reasons cited more frequently were that the agent made a decision category i),<sup>14</sup> acted on his own (category ii),<sup>15</sup> and was not caused to act by external causes (category iii).<sup>16</sup> All three reasons can be figured out from the sole vignette without having to compare it to a 'normal' case without counterfactual intervener. This suggests that participants' widely shared intuition that Mr. Jones can be morally responsible even though he could not have done otherwise is elicited by features that are internal to the actual case, and not by a comparison to some other 'normal' case.

Of course, one might doubt the reliability of our participants' answers, and insist on the fact that introspection is not always reliable. However, the participants' answers are in touch with plausible psychological accounts of how people attribute moral responsibility for a given outcome. For example, Cushman (2008) proposes that people attribute moral responsibility mainly by checking whether the agent (i) caused a given outcome and (ii) caused it on the basis of his intention or desire to cause it. Similarly, Cova et al. (submitted) claim that people generally attribute moral responsibility to agents by following a very simple "valence-matching heuristic" according to which an agent is morally responsible for a given outcome if (i) the agent caused this outcome and (ii) caused it on the basis of motives that have the same moral valence as the outcome. As a result, there is no compelling reason to doubt that people can have (and actually have) the intuition that agents in FSCs are morally responsible without having to compare these cases to cases without counterfactual interveners.

Finally, one might worry that the participants' answers to Question 3 (the *Responsibility* question) do not truly reflect these participants' "intuitions", but rather an elaborated answer due to thorough reasoning. However, this worry can be assuaged by participants' answers to Question 6: when asked whether they had to think a lot to answer Question 3 or it was easy, 156 participants out of 179 (87 %) answered that it was in fact easy.<sup>17</sup> Moreover, the average

<sup>12</sup> Only category (iv) was predicted *a priori*. Other categories were constructed from participants' answers to the pilot study I referred to in the previous footnotes.

<sup>13</sup> Cohen's kappa=.74.

<sup>14</sup> For example, one participant answered that: "Because Jones decided to steal the car, he is responsible."

<sup>15</sup> For example, one participant answered that: "Mr. Jones stole the car on his own. The decision-making device never was activated." (This particular answer also fell into category iii).

<sup>16</sup> For example, one participant answered that: "He did it without the device changing his impulses".

<sup>17</sup> One can trust participants' introspection on this question, since participants who answered that it was easy spent less time on the questionnaire ( $M=130.3$  s) than participants who answered they had to think a lot ( $M=168.4$  s) and that this difference was statistically significant (Welch t-test:  $N=179$ ,  $t=-2.9$ ,  $df=25.2$ ,  $p<0.01$ ).

t2.1 **Table 2** Categories of explanations in participants' answers for Question 4

t2.2		(i)	(ii)	(iii)	(iv)	(v)
t2.3		Choice	On his own	Not forced	Comparison 1	Comparison 2
t2.4	<i>Coder EP</i>	153 (94 %)	146 (90 %)	89 (55 %)	2 (1 %)	36 (22 %)
t2.5	<i>Coder VS</i>	131 (80 %)	110 (67 %)	88 (54 %)	1 (0 %)	26 (16 %)
t2.6	<i>Trials on which both coders agreed</i>	133 out of 144 (92 %)	110 out of 125 (88 %)	76 out of 135 (56 %)	0 out of 160 (0 %)	22 out of 147 (15 %)

time participants spent on the questionnaire was 135.2 s. Given that this time includes reading the vignette and answering all six questions (including an open-ended one), it is safe to think that participants' answers to Question 3 are more likely to reflect participants' intuitive response to the vignette rather than the conclusion of careful reasoning.

Thus, it is reasonable to suppose that FSCs have the property of directly eliciting the intuition that the agent is morally responsible even if he could not have done otherwise. This allows for their use in what I will call the *direct* Frankfurt counter-argument (henceforth: DFA):

1. The agent is morally responsible in the *Frankfurt* case.
2. However, due to the presence of the counterfactual intervener, the agent could not have done otherwise in the *Frankfurt* case.
3. Therefore, the PAP is false, since the agent in the *Frankfurt* case is morally responsible even though he could not have done otherwise.

I should note that DFA is not an invention of my own, although it is true that, in his original 1969 paper, Frankfurt used his FSCs in support of what I have called IFA. For example, commenting on a FSC in which the agent is called Jones and in which the counterfactual intervener is called Black, Frankfurt writes:

suppose that Black never has to show his hand because [Jones], for reasons of his own, decides to perform and does perform the very action Black wants him to perform. In that case, it seems clear, [Jones] will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he did. It would be quite unreasonable to excuse [Jones] for his action, or to withhold the praise to which it would normally entitle him, on the basis of the fact that he could not have done otherwise. This fact played no role at all in leading him to act as he did. He would have acted the same if it had not been a fact. Indeed, everything happened just as it would have happened without Black's presence in the situation and without his readiness to intrude into it. (p.836)

Here, we have a perfect example of what Levy attacks; the argument is grounded in a comparison between the FSC and its 'normal' counterpart. However, Frankfurt's initial intentions notwithstanding, FSCs are not always used in this way in the philosophical literature. For example, in chapter 8 of his *Contemporary Introduction to Free Will*, Robert Kane (2005) introduces the case of Jones and Black without reference to its 'normal' counterpart and comments it this way:

The point of the example is this: Jones *cannot do otherwise* because Black will not let him. But Jones might decide *on his own* to do what Black wants, in which case Black would not intervene. Frankfurt's claim is that if Jones does act on his own and Black

does not intervene, then Jones would be responsible for what he did even though Jones could not have done otherwise. For if Jones did act on his own without Black intervening, Jones would have done so from his own motives and for his own reasons and no one would have interfered with his choice. The principle which says that moral responsibility requires alternative possibilities (PAP) would then be false: Jones would have acted responsibly though he could not in fact have done otherwise (because Black would not have let him). (p.84)

Here, there is no comparison between the FSC and another case. The conclusion that the agent is morally responsible even if he could not have done otherwise is drawn from features of the FSC itself. However, it is not clear that Kane's point is that it is directly intuitive that the agent is morally responsible; there seems rather to be an argument according to which the agent is morally responsible because he acts on his own and nothing interferes with his choice. But such an argument seems only to beg the question against the PAP, for its acceptance requires that we agreed that all that matters is that the agent acts on his own. The argument is a better one if we understand it as claiming (i) that we have the intuition in such cases that the agent is morally responsible and (ii) that we have this intuition because the agent acts on his own. Understood this way, the argument no longer begs the question against the PAP but makes use of our intuitions to falsify the PAP and argues that all that matters is that the agent acts on his own.

A clearer and less controversial use of DFA can be found in a paper recently published in this journal. In this paper, Di Nucci (2010), commenting on Jones and Black's case, writes:

Frankfurt comes up with a scenario in which the agent is intuitively morally responsible for what he does, even though the agent could not have done otherwise. Therefore the Principle is false. (p.208)

Here, we have a paradigmatic example of how FSC can be used in a *direct* Frankfurt counter-example: PAP fails because it is *intuitive* that agents in FSCs are morally responsible.

Though it is not how Frankfurt intended them to be used, my contention is that both kinds of counter-arguments can be found in the literature and in philosophy courses, and that both are valid arguments.

Now, one could be reluctant to consider DFA as a compelling argument. One reason could be that it seems such a *crude* argument, relying only on appeal to intuitions. However, there are well-known examples in the philosophical literature of counter-arguments relying mainly on intuitions about particular cases. We just have to think about Gettier cases in epistemology, the Knowledge Argument in philosophy of mind, or Nozick's Experience Machine in ethics. If those are respectable philosophical arguments, then there is no reason to take DFA seriously.

However, one could argue that there is a significant difference between these cases and DFA: while the justified true-belief conception of knowledge, physicalism and hedonism are only philosophical theses, the PAP is also *intuitive* (or *self-evident*). It is not a standard thesis but a *prima facie* principle that has long been taken as a starting point, and not as a conclusion one should reach. Thus, using the direct Frankfurt counter-argument against the PAP amounts to just clashing two intuitions against each other: nothing compelling can come out of such a clash, since both intuitions cancel each other.

There is some truth to this objection. Nevertheless, we have several ways of solving this difficulty. A first move could be to point that IFA also involved an appeal to intuitions: how else are we supposed to know that the agent is responsible in the 'normal' case? Thus, this objection doesn't seem to make DFA *weaker* than IFA. However, one could argue that this

intuition is widely shared, while the direct intuitions about Frankfurt cases are more controversial, and thus not *strong* or *evident* enough to overcome the PAP's intuitiveness.

In answer, we can attack the PAP's alleged intuitiveness. If the PAP is so self-evident and intuitive, how is it possible that philosophers have long debated how it should be interpreted? Clearly, the PAP according to the compatibilist interpretation is not the same than the PAP according to the incompatibilist interpretation; although their wording can be the same, this resemblance is not enough to make them two instances of the same principle. So it seems that there is no single PAP likely to be evident to all, for compatibilists and incompatibilists endorsed at least two different PAPs. If we now take into account the diversity of compatibilist interpretations of the PAP, and the multiplicity of alternate principles (like, for example, the Principle of Possible Prevention<sup>18</sup>) that have been proposed by incompatibilists, it seems that there is no single PAP that is intuitive enough to resist the widely shared intuition elicited by FSCs. At best, we can say that there is "a widely accepted moral intuition" (to quote Widerker 2002) that something like the PAP is true. But this is an affirmation that is perfectly compatible with the falsity of the PAP and our intuitions in FSCs.

This leads us to our third possible answer: that when two intuitions clash, we can argue in favor of the first by advancing an error-theory accounting for the second. And it turns out that Frankfurt's original paper proposes a compelling error-theory for the intuition that the PAP is true. According to Frankfurt, the PAP should be replaced by the principle according to which "a person is not morally responsible for what he has done if he did it only because he could not have done otherwise" (p.838). Frankfurt comes to this conclusion by reflecting on our use and standards of acceptance for excuses such as "I could not have done otherwise". For him, we would reject such an excuse if we discovered that its enunciator did not act the way he did *only because* he could not have done otherwise. On the contrary, according to Frankfurt, "we understand the person who offers the excuse to mean that he did what he did *only because* he was unable to do otherwise, or *only because* he had to do it. And we understand him to mean, more particularly, that when he did what he did, it was not because that was what he really wanted to do" (p.838). From this suggestion, we can sketch the following error-theory: the real intuitive principle is that we should excuse people for actions they performed *only because* they were unable to do otherwise, and this is what is conveyed by the current expression "I could not have done otherwise". However, the fact that this expression means more than is apparent from its surface grammar led philosophers to think that the intuitive principle was the PAP. Frankfurt's principle and the PAP being close enough to allow such a confusion, we now have an error-theory allowing us to keep both intuitions that "something like the PAP" must be true and that agent in FSCs are morally responsible for their actions *even if* they could not have done otherwise.<sup>19</sup>

Thus, I conclude that DFA is as valid and compelling a philosophical argument as IFA. However, as DFA does not rely on a comparison between cases, it requires no premise according to which the mere presence of a counterfactual intervener has no effect. Thus, it is not vulnerable to Levy's objection. More precisely, there are two possibilities: either the mere presence of a counterfactual intervener makes no difference, in which case IFA succeeds, or it makes a difference, and DFA leads us to the conclusion that the ability the

<sup>18</sup> See van Inwagen (1983).

<sup>19</sup> In his paper, Frankfurt (1969) also sketches an alternative error-theory, according to which the PAP would derive some of his plausibility from a misleading psychological association with what he calls the 'doctrine of coercion' (according to which moral responsibility is excluded by coercion). Frankfurt then proceeds to show that (i) either the doctrine of coercion is wrong (ii) either the truth of the PAP cannot be inferred from its truth.

counterfactual intervener made disappear (such as the ability to do otherwise) is not relevant to moral responsibility. 543  
 544

**3 Objections and Replies** 545

What moves are now open to Levy? The obvious one is to attack DFA, for it will leave him just IFA to contend with, and I have presupposed that his argument succeeds against IFA. In order to defeat DFA, Levy could use the traditional criticism of FSCs—the flickers of freedom objection or the dilemma defense. However, if these arguments work against DFA, they work as well against IFA. This means that whoever is already convinced by these arguments to reject DFA will also be convinced to reject IFA, making Levy’s argument superfluous. In other words, for his argument to work as an argument against both DFA and IFA, Levy should admit that there is already an argument that works against both DFA and IFA, thus endorsing the conclusion in the premises, and making the whole argument circular. Of course, this would not prevent Levy’s argument from working as a modest refutation of the sole IFA. But my aim here is to reject Levy’s contention that his argument is strong enough to change on its own the face of the debate about alternative possibilities.<sup>20</sup>

Levy can thus be tempted to reject appeal to intuitions, based (for example) on their unreliability. This argument runs the risk of also affecting IFA, making once again Levy’s original argument superfluous. But Levy can say that the ‘normal’ cases used IFA do not display the weird features of FSCs (such as neural devices and counterfactual intervener) and look ordinary enough for our intuitions to be warranted in those cases – or he could also say that it is given ‘by construction’ that agents are responsible in ‘normal’ cases. This leaves us with the following line of answer: direct intuitions about cases as weird as FSCs are too unreliable, which makes IFA the only way to undermine PAP by use of FSCs, but IFA is itself undermined by Levy’s argument.

The problem with this objection is that it backfires on Levy’s original argument: FECs such as *Phobia*, which Levy uses for his demonstration, are from Levy’s own admission very close to FSCs. Then if intuitions in FSCs cannot be trusted, it seems that they cannot be trusted in FECs either, ruining Levy’s initial demonstration.<sup>21</sup> It seems that, to kill the DFA, Levy must spare IFA.

A first possible move would be to grant that this kind of skepticism indeed undermines Levy’s original argument, but not enough to warrant our intuitions about IFA, so that the main goal of rescuing the PAP is ensured. According to this solution, intuitions about FSCs and FECs cannot be trusted, which is why DFA but also Levy’s own original argument fail. However, the mere presentation of FECs and of our intuitions about them is enough to show that it *could be*

<sup>20</sup> For example, Levy (2008) concludes his paper this way: “For the time being, FSCs should not be seen as having broken the dialectical stalemate that characterized the free will debate before they burst onto the scene” (p.239). Similarly, in a later paper on the same argument (Levy 2012), he writes: “I argued that [Frankfurt-style cases] are not sound. Or rather, I argued that we are not entitled to the conclusion to which they seem to lead us: that agents can be free or morally responsible, despite lacking alternative possibilities” (p.607). Thus, Levy thinks that his argument leads to the conclusion that FSCs cannot be used to undermine PAP. It is this conclusion I reject here, and not the most modest conclusion that his argument sheds doubt on one of the ways FSCs can be used to undermine PAP.

<sup>21</sup> Note that this is not because (direct) intuitions about FSCs and FECs clash and cannot be true at the same time. If we grant that the presence of the neuroscientist make Jillian gain an ability in FECs, this shows only that the mere presence of a counterfactual intervener can change an agent’s ability. This is perfectly compatible with the claim that (i) agent in FSCs do not lose abilities or (ii) that they lose abilities but are still morally responsible.

that the mere presence of a counterfactual intervener is enough to modify an agent's abilities, and this mere possibility is enough to cast doubt on IFA as well. We are then left with no possibility of knowing if DFA and IFA's premises are true, and without means to reject the PAP.

Nevertheless, this won't do. Such a move presupposes that the mere possibility that a counter-argument can be true is enough to cast doubt on an argument (here the mere possibility that Levy's argument can be true is enough to cast doubt on IFA). But, in this dialectical situation, one has not proved that either DFA or IFA is false: he has just cast doubt on their truth. Thus, it is still possible that these arguments can be true. And, if we follow the same epistemic rules, this means we should also doubt the PAP, because DFA and IFA are counter-examples to PAP. We are then led to a wide skepticism that engulfs both the PAP and Frankfurt counter-arguments.<sup>22</sup>

One might say that, even in such a case, Levy would have achieved his goal of returning to the 'pre-Frankfurt' dialectical stalemate. But this is not that simple: the elimination of the PAP benefits to compatibilism, since compatibilism is, from a logical point of view, the default position. As William Lycan explains, "a theorist who maintains of something that is not obviously impossible that nonetheless that thing is impossible owes us an argument" (Lycan 2003; see also Nahmias et al. 2006). If the PAP cannot longer be used, then the stalemate is broken and compatibilists victorious (though one may think there is nothing glorious in such a victory by default).

One might object that, even if compatibilism is the default position if we limit ourselves to a logical point of view, it is no longer the default position once we take our intuitions on the matter into account. Indeed, many authors have claimed that incompatibilism is people's default view and that people have to be argued into compatibilism (for a sample of quotes illustrating this view, see Nahmias et al. 2006). However, this claim does not fare so well when confronted to experience. Recently, philosophers have studied empirically people's intuitions about free will, moral responsibility, and determinism. Of course, their results have not been straightforward and have given birth to opposite interpretations. However, two main interpretations have come to dominate the debate (for a review, see Cova 2011; Cova & Kitano, in revision). On the first interpretation, people have conflicting intuitions and have compatibilist intuitions in some contexts and incompatibilist intuitions in others (see Nichols and Knobe 2007; Weigel 2012). On the second interpretation, people's default view is compatibilist, and apparent incompatibilist intuitions can be dismissed as comprehension errors (Nahmias and Murray 2010; Murray and Nahmias in press; Cova et al. 2012). In both cases, our claim that compatibilism constitutes a default position is warranted. If the latter interpretation is right, then compatibilism is the default position at both the logical and intuitive levels. If the former interpretation is right, then there is no real intuitive default point, and we are brought back to the logical level, at which compatibilism is the default position. Therefore, casting doubt on PAP seems to give the dialectical advantage to compatibilism.<sup>23</sup>

What Levy needs then, if he wants to reach the conclusion that we are stuck again in a dialectical stalemate, is a way to cast doubt on intuitions about FSCs without casting doubt

<sup>22</sup> Of course, as a reviewer for this journal rightly points out, this does not mean that choosing this solution commits one to doubt any intuition about moral responsibility (such as intuitions about 'regular' and mundane cases). All I want to stress is that if one thinks that Levy's argument is still strong enough to undermine IFA when intuitions about FECs and FSCs are unreliable, then one must conclude that DFA and IFA are also still strong enough to undermine PAP in these conditions.

<sup>23</sup> Some might not agree with my claim that current results in experimental philosophy cannot be interpreted in a way that makes them compatible with the hypothesis that people are naturally drawn to incompatibilism. However, even if they could, it would still be true that one cannot say that current results warrant the thesis that people's default view is incompatibilism. Thus, at best, one could say that no conclusion about people's default position has been established. Once again, in absence of a clear intuitive default position, this would bring us back to the logical level, at which compatibilism is the default position.

on PAP or Levy's own argument. One way to do that is, as we saw, to propose some kind of error-theory. And that is what Levy seems to be doing; he claims that intuitions about FSCs are unreliable because they rely on the hypothesis that a mere counterfactual intervener has no effect on the agent's responsibility-grounding abilities, and that this belief can be traced back to a form of internalism about agents' abilities.

I agree with Levy that, as soon as we grant that the agent can gain genuine responsibility-grounding abilities in FECs, we have here a compelling story that should lead us to doubt premise 2 of IFA, that is the premise according to which there is no significant difference between the 'normal' and the Frankfurt case as far as the agent's abilities are concerned. However, this story leaves unaffected the intuition on which DFA is grounded. As we saw, there is no reason to think that this intuition is based on such beliefs. In fact, there are even reasons to think that this intuition is perfectly sensitive to the changes in the agent's abilities, for it is the same kind of intuitions that Levy uses in FECs to track change in agent's abilities. Once again, if Levy thinks that intuitions in FSCs are unreliable because they cannot track changes in the agent's abilities, then the same lesson must be drawn for intuitions about FECs, depriving Levy's argument of its best evidence.

The last way open to Levy is then to find a way of discrediting direct intuitions about FSCs while keeping intuitions about FECs reliable and trustworthy. I confess that I am unable to think of an argument for treating FECs and FSCs differently that does not straightforwardly beg the question. A (not so much) plausible solution might be to argue that our intuitions track changes in abilities when abilities are gained (like, supposedly, in FECs) but not when intuitions are lost (like, supposedly, in FSCs). However, it is easy to design a case in which our intuitions seem (by Levy's standards) to be sensitive to a loss in responsibility-grounding abilities. You just have to take *Phobia* and reverse it in the following way: Jillian has the swimming abilities required to save the child but no desire to save him. However, unbeknownst to her, an evil neuroscientist planted in her brain a chip that, if activated, can cause a terrible fear of water. Had Jillian tried to save the child, the neuroscientist would have activated the chip, and Jillian could not have saved the child. In this case, I think that we have the intuition that, although Jillian deserves blame for not trying to save the child, she is not responsible for the child's death, since she could not have saved him (she did not have the *possibility* to save him). If I am right, then our direct intuitions about FSCs and FECs are, by Levy's standards, sensitive to both gains and losses in responsibility-grounding abilities.

To sum up, I have distinguished two ways FSCs can be used to falsify PAP: the *indirect* and the *direct* Frankfurt counter-examples. While Levy's argument might—at best—save the PAP from the indirect counter-argument, I argue that it cannot refute the direct counter-argument, unless we adopt some wide skepticism that will take down the PAP as well. In conclusion, we are not back to the dialectical stalemate the PAP had produced and philosophical reflection about moral responsibility and determinism can move on.

**Acknowledgments** For their useful and challenging comments on previous versions of this paper, I would like to thank Hichem Naar and two anonymous reviewers for this journal. I also thank Eva Pool and Vanessa Sennwald for coding the experimental data. Finally, I thank the Swiss Centre for Affective Sciences and the University of Geneva, where this research was carried.

**References**

Clarke R (2011) Responsibility, mechanisms, and capacities. *Mod Sch* 88:161–169  
 Cova F (2011) *Qu'en pensez-vous ? Une introduction à la philosophie expérimentale*. Germina, Paris

Cova F, Dupoux E, Jacob P (submitted) The more you deserve blame, the more you deserve blame: The valence matching heuristic	663 <sup>Q3</sup> 664
Cova F, Kitano Y (in revision) Experimental philosophy and the compatibility of moral responsibility and determinism: a survey	665 <sup>Q4/Q5</sup> 666
Cova F, Bertoux M, Bourgeois-Gironde S, Dubois B (2012) Judgments about moral responsibility and determinism in patients with behavioural variant of frontotemporal dementia: still compatibilists. <i>Conscious Cogn</i> 21(2):851–864	667 668 669
Cushman F (2008) Crime and punishment: distinguishing the roles of causal and intentional analyses in moral judgment. <i>Cognition</i> 108:353–380	670 671
Di Nucci E (2010) Refuting a frankfurtian objection to frankfurt-type counterexamples. <i>Ethical Theory Moral Pract</i> 13(2):207–213	672 673
Fischer JM (1994) <i>The metaphysics of free will</i> . Blackwell Publishers, Oxford	674
Fischer JM (2006) <i>My way</i> . Oxford University Press, New York	675
Frankfurt HG (1969) Alternate possibilities and moral responsibility. <i>J Philos</i> 66:829–839	676
Ginet C (1996) In defense of the principle of alternate possibilities: why i don't find Frankfurt's argument convincing. <i>Philos Perspect</i> 10:403–417	677 678
Haji I, McKenna M (2011) Disenabling levy's frankfurt-style enabling cases. <i>Pac Philos Q</i> 92:400–414	679
Kane R (2005) <i>A contemporary introduction to free will</i> . Oxford University Press, New York	680
Levy N (2008) Counterfactual intervention and agents' capacities. <i>J Philos</i> 105:223–239	681
Levy N (2012) Capacities and counterfactuals: a reply to Haji and McKenna. <i>Dialectica</i> 66:607–620	682
Lycan W (2003) Free will and the burden of proof. In: O'Hear A (ed) <i>Proceedings of the royal institute of philosophy for 2001–02</i> . Cambridge University Press, Cambridge, pp 107–122	683 684
Miller J, Feltz A (2011) Frankfurt and the folk: an experimental investigation of Frankfurt-style cases. <i>Conscious Cogn</i> 20:401–414	685 686
Murray D, Nahmias E (in press) Explaining away incompatibilist intuitions. <i>Philos Phenomenol Res</i>	687 <sup>Q6</sup>
Nahmias E, Murray D (2010) Experimental philosophy on free will: An error theory for incompatibilist intuitions. In: Aguilar J, Buckareff A, Frankish K (eds) <i>New waves in philosophy of action</i> . Palgrave-Macmillan, New York	688 689 690
Nahmias E, Morris S, Nadelhoffer T, Turner J (2006) Is incompatibilism intuitive? <i>Philos Phenomenol Res</i> 73:28–53	691 692
Nichols S, Knobe J (2007) Moral responsibility and determinism: the cognitive science of folk intuitions. <i>Noûs</i> 41:663–685	693 694
Pereboom D (1995) <i>Determinism al Dente</i> . <i>Noûs</i> 29:21–45	695
Pereboom D (2001) <i>Living without free will</i> . Cambridge University Press, New York	696
van Inwagen P (1983) <i>An essay on free will</i> . Clarendon, Oxford	697
Weigel C (2012) Experimental evidence for free will revisionism. <i>Philos Explor</i> 16:31–43	698
Widerker D (2002) Responsibility and frankfurt-type examples. In: Kane R (ed) <i>The Oxford handbook of free will</i> . Oxford University Press, Oxford, pp 323–334	699 700 701

## AUTHOR QUERIES

### **AUTHOR PLEASE ANSWER ALL QUERIES.**

- Q1. Please check the suggested running page title if appropriate. Otherwise, please provide short running title with maximum of 65 characters including spaces.
- Q2. Missing citation for Table 1 was inserted here. Please check if appropriate. Otherwise, please provide citation. Note that the order of main citations of tables in the text must be sequential.
- Q3. Please provide updated information in reference (Cova et al. submitted).
- Q4. Cova F (2999b) was not cited anywhere in the text. Please provide a citation. Alternatively, delete the item from the list.
- Q5. Please provide updated information in reference (Cova and Kitano in revision).
- Q6. Please provide updated information. If possible, please advise the DOI number if this reference is still for publication or in press (Murray and Nahmias in press).