

ON THE VERY IDEA OF A ROBUST ALTERNATIVE

CARLOS J. MOYA
Universidad de Valencia
Carlos.Moya@uv.es

SUMMARY: According to the Principle of Alternative Possibilities (PAP), an agent is morally responsible for an action of hers only if she could have done otherwise. The notion of a robust alternative plays a prominent role in recent attacks on PAP based on so-called Frankfurt cases. In this paper I defend the truth of PAP for blameworthy actions against Frankfurt cases recently proposed by Derk Pereboom and David Widerker. My defence rests on some intuitively plausible principles that yield a new understanding of the concept of a robust alternative. I will leave aside whether PAP also holds for praiseworthy actions.

KEY WORDS: Principle of Alternative Possibilities, Frankfurt cases, blameworthiness, Pereboom, Widerker

RESUMEN: Según el Principio de Posibilidades Alternativas (PPA), un agente es moralmente responsable de una acción sólo si hubiera podido actuar de otro modo. La noción de alternativa robusta desempeña un papel prominente en ataques recientes al PPA basados en los llamados casos Frankfurt. En este artículo defiendo el PPA para la culpabilidad moral frente a casos Frankfurt propuestos recientemente por Derk Pereboom y David Widerker. Mi defensa descansa en algunos principios intuitivamente plausibles que dan lugar a una comprensión nueva del concepto de alternativa robusta. No trataré la cuestión de la verdad del PPA para acciones moralmente laudables.

PALABRAS CLAVE: Principio de posibilidades alternativas, casos Frankfurt, culpabilidad, Pereboom, Widerker

1. *Introduction: Some Background*

The general view that alternative possibilities are necessary for moral responsibility (MR for short, in what follows) finds a particular expression in the so-called Principle of Alternative Possibilities (PAP, for short). According to PAP, an agent is morally responsible for what she has decided and done only if she could have decided and done otherwise, or at least if she could have avoided deciding and acting as she did. We shall restrict our concern mainly to decisions, with occasional reference to overt actions. In addition, we will deal with PAP in connection with moral blameworthiness only, leaving aside praiseworthiness. PAP was once taken to be almost self-evidently true, but is nowadays under strong pressure. In fact, many

philosophers think it is false. Their main reasons for this position have to do with so-called Frankfurt cases.¹

Frankfurt cases are supposed to be conceptually possible situations in which an agent, on her own and for her own reasons, makes a certain decision which, owing to circumstances of which she is fully unaware, is the only one she can actually make; now, the circumstances that make an alternative decision impossible do not cause or influence in any way her actual decision; she makes this decision spontaneously, with no hindrance or coercion and on the basis of suitable reasons; in situations with these features, it seems that the agent is morally responsible for her decision, though she could not have decided otherwise. If this is so, then PAP is false.

Part of what fuels the intuition of the agent's MR in Frankfurt cases is that, given the causal insignificance of the circumstances that rule out an alternative decision, the actual decision is the one the agent would have made anyway, even if those circumstances had been absent and she *could* have decided otherwise. So, these cases are supposed to show that having access to alternative decisions is irrelevant to an agent's MR for the decision she actually makes. What is important for such MR is the actual causal history of the decision and whether this causal history is of the right sort, so that it does not contain coercion or any other factors that are commonly taken to diminish or rule out MR; it does not matter whether alternative decisions were available or whether the actual decision might have had a different causal history. Against this "actual history" or "actual sequence" view of MR, PAP suggests instead the view that what an agent *can* do or *could* have done is also relevant to the MR she bears for what she does. I will try to show that the latter view is correct.

Original Frankfurt cases, designed by Frankfurt himself (cf. Frankfurt 1969), feature an agent who decides and does on her own something which, unknown to her, she would be caused to decide and do anyway by an alien factor if she were to show some sign that she was not going to decide and do it. The following is a case of this sort. In a situation with morally significant profiles, Betty is deliberating about whether to lie or to tell the truth to a friend of hers concerning an important matter. Black, a nefarious neurosurgeon, wants Betty to lie and, unbeknownst to her, has implanted in her brain a device that allows him to follow Betty's deliberation; by means of this de-

¹ They take this name from Harry Frankfurt's pioneering article "Alternate Possibilities and Moral Responsibility" (Frankfurt 1969), where he first designs such cases.

explanatory irrelevance would be to contend that Betty is morally responsible for deciding to lie because she might not have decided this as a consequence of her possible sudden death. This explanation of her MR is obviously wrong. A robust alternative would be her free decision to tell the truth, but, given the features of the situation she finds herself in, this decision is not available to her.

The flicker theorist may point instead to the aforementioned alternative of Betty's showing the sign that would have prompted Black's intervention. This sign might well be an inclination of hers towards telling the truth, or her paying attention to moral reasons in favour of this alternative. Tiny as these episodes may be, they are plausibly taken to be under the agent's control. Fischer's response to this move was to point out that Frankfurt cases can be designed where the sign for Black's intervention is a mere happening, beyond the agent's control, such as a blush or a certain neurological pattern in the agent's brain (cf. Fischer 1994, p. 144). Assuming that, in the preceding example, it was a sign of this sort that alerted Black of Betty's future decision to tell the truth, it would be ludicrous to claim that Betty's was morally responsible for lying because she could have blushed or shown a certain neurological pattern. These alternative possibilities would have been mere happenings, fully behind Betty's control.

At this point in the dialectic, however, some authors (Kane 1985; Ginet 1996; and especially Widerker 1995) have developed an important argument in favour of PAP, known as the "dilemma defence". It can be formulated as follows. Think of the sign that Black uses *to remain inactive*, say a blush of Betty's at a certain moment, t_1 , of her deliberation process, prior to her decision to lie at t_2 . Concerning this sign, the Frankfurt theorist has to confront the following dilemma: either this sign at t_1 is (or is associated with a condition that is) causally sufficient for Betty's decision to lie at t_2 or it is not. If it is, then this decision is causally determined; but incompatibilists, who hold that MR and causal determinism cannot coexist, will not accept Betty's MR. If it is not —so that it is only a reliable, but not infallible, symptom of Betty's later decision— she may be morally responsible for this decision, but then there is no clear reason to think that she could not have decided (and done) otherwise. Either way, the proponents contend, PAP remains safe.

The dilemma defence of PAP has put strong pressure on the construction of plausible Frankfurt cases. In the face of it, most Frankfurt theorists, understandably enough, have designed cases that do not assume determinism. This paper will be concerned with cases

of this sort. Before going to them, however, it is convenient to say something about the deterministic horn of the dilemma, for some Frankfurt theorists do not think that assuming determinism in Frankfurt cases begs the question against incompatibilists. They include, for instance, Fischer (1999, 2010) and Haji and McKenna (2004, 2006).

Haji and McKenna (2004) contend that so-called “leeway” incompatibilists, who base their incompatibilism on the claim that determinism rules out alternative possibilities, are not entitled to hold that deterministic Frankfurt cases beg the question against them, for the incompatibility between causal determinism and MR does not follow only from the fact that determinism excludes alternatives, but from it plus PAP. But PAP is precisely what is at issue in Frankfurt cases. Only so-called “source” incompatibilists, who hold that determinism rules out MR because it precludes agents from being the true source of their decisions and actions, can reject deterministic Frankfurt cases legitimately. But even they should accept one central moral of Frankfurt cases, namely that if the agent in such cases is not responsible for her decision, this is not because she lacks alternatives.² So, PAP is undermined by such cases anyway, even if they do not prove formally that it is false.

What can a dilemma defender respond to this objection to the deterministic horn? I would think that the following remarks constitute a plausible rejoinder.

The objection starts from a clear-cut distinction between leeway- and source-incompatibilism. But this distinction has been mainly a result of reflection on Frankfurt cases, whose success is precisely at stake in the debate. In rejecting this success, dilemma defenders can also reject the indicated clear-cut distinction and contend, for example, that true sourcehood involves alternative possibilities. They can plausibly hold that an agent cannot be considered as the true source and author of her decision, and so able to make a difference to the course of events, if this decision was the only one that she could possibly have made, as presumably happens if determinism is true. Now, if sourcehood and alternatives are intermingled in this way, an incompatibilist can reject the agent’s MR in deterministic Frankfurt cases in a legitimate way.

In addition, it seems to me that deterministic Frankfurt cases violate a central internal condition of successful Frankfurt cases, one

² I am grateful to Ishtiyaque Haji for helping me with this point in a written comment on a previous version of this paper.

that Frankfurt himself endorses, namely that, in them, the circumstances that make it impossible for the agent to decide otherwise do not cause or bring it about in any way that she decides as she actually does (cf. Frankfurt 1969, p. 9). Now suppose that the agent's actual decision is ensured in that it is the effect of a causally deterministic chain that starts with an event beyond the agent's control. It would seem that a case like this violates the aforementioned condition, for in it the circumstances that rule out alternatives *also* contribute to causing the decision, even if the agent is not aware of this fact. For this reason, if Frankfurt cases are to elicit a clear intuition about the agent's MR, it is important that alternatives are excluded only by means of a purely counterfactual intervener.³

Finally, as a general remark, consider that PAP relates only two notions, namely MR and alternative possibilities. As such, it does not mention determinism. And I think that criticizing PAP by means of examples that include determinism is likely to distort the intuitions they raise so as to make them unsteady and inconclusive.

For all these reasons, I tend to think that it is good advice for opponents to PAP to embrace the indeterministic horn of the dilemma. And in fact most of them have chosen this option. The task for these theorists is to design cases that meet the following requirements. First, causal determination of the agent's decision is not assumed to hold in the "actual" sequence. Second, the agent has no access to robust alternatives, such as alternative decisions. Finally, the agent is clearly morally responsible or, more exactly, morally blameworthy, for her decision.⁴ Widerker's (2009) and Pereboom's (2001, 2003, 2009, forthcoming) recent examples are supposed to meet these re-

³ Some recent Frankfurt cases, such as Mele and Robb 2003, and Haji 2010, assume that two causal chains are present in the actual sequence, one deterministic and the other indeterministic. The latter corresponds to the agent's deliberation and it is the only one that causes her decision. I think that in cases of this sort the deterministic chain plays the role of counterfactual interveners in classical Frankfurt cases, since it is only a failsafe device that never causes the decision. Depending on the particular features of the case, a PAP defender may either (a) accept the agent's MR in such cases but contend that, in them, the agent has robust alternatives, or (b) reject her MR on the basis that the agent's mechanism of deliberation and decision is not responsive to reasons. The arguments for these claims will be roughly the same as those we will be developing below against Pereboom's and Widerker's examples.

⁴ As I said above, I restrict myself to a defence of PAP for blameworthiness; I leave aside the question of praiseworthiness; so this paper is consistent with an asymmetrical approach to PAP, according to which alternatives are necessary for being blameworthy, but not for being praiseworthy, for what one does. In fact, I have tentatively defended an asymmetrical view in Moya 2010.

quirements. I will examine them below. My contention will be that these examples violate some of these conditions. We start with Pereboom's example.

2. *Pereboom's Post-Dilemma Example*

This is Pereboom's example:

Tax Evasion (2): Joe is considering claiming a tax deduction for the registration fee that he paid when he bought a house. He knows that claiming this deduction is illegal, but that he probably won't be caught, and that if he were, he could convincingly plead ignorance. Suppose he has a strong but not always overriding desire to advance his self-interest regardless of its cost to others and even if it involves illegal activity. In addition, the only way that in this situation he could fail to choose to evade taxes is for moral reasons, of which he is aware. He could not, for example, [fail to] choose to evade taxes for no reason or simply on a whim. Moreover, it is causally necessary for his failing to choose to evade taxes in this situation that he attain a certain level of attentiveness to moral reasons. Joe can secure this level of attentiveness voluntarily. However, his attaining this level of attentiveness is not causally sufficient for his failing to choose to evade taxes. If he were to attain this level of attentiveness, he could, exercising his libertarian free will, either choose to evade taxes or refrain from so choosing (without the intervener's device in place). However, to ensure that he will choose to evade taxes, a neuroscientist has, unbeknownst to Joe, implanted a device in his brain, which, were it to sense the requisite level of attentiveness, would electronically stimulate the right neural centers so as to inevitably result in his making this choice. As it happens, Joe does not attain this level of attentiveness to his moral reasons, and he chooses to evade taxes on his own, while the device remains idle. (Pereboom 2009, p. 113; cf. 2001, and 2003, p. 193)

As we see, the actual sequence is explicitly assumed to be indeterministic and Joe is even depicted by Pereboom as having a libertarian free will. A necessary condition for Joe's failing to decide to evade taxes is a voluntary mental act, namely to reach a certain level of attentiveness to moral reasons against evading taxes. Joe could have performed that mental act but did not, and decided on his own to

evade taxes. Joe, however, could not have failed to make that decision, for, had he reached the required level of attention to moral reasons, the device in his brain would have been activated and would have causally induced in him the decision to evade taxes anyway.

Pereboom's example raises some worries concerning the deterministic horn of the aforementioned dilemma, for, if Joe's attaining a level of attentiveness to moral reasons is causally necessary for him to fail to decide to evade taxes, his *not* attaining that level, as is actually the case, is causally sufficient for him *not to fail* to decide to evade taxes, that is, for his actual decision to evade taxes, which would be then causally determined. I think that Pereboom can meet this objection by holding that reaching the required level of attentiveness is under Joe's voluntary control until the very moment of his choice, so that this remains causally undetermined (cf. Moya 2006, p. 57, and Pereboom 2003, p. 195). But let me concentrate on our main concern in this paper, namely, the agent's access to robust alternatives.

A PAP defender may argue that Joe's attaining a certain level of attentiveness to moral reasons is not a mere happening beyond Joe's control, but an act that he could have freely and voluntarily performed, as Pereboom himself acknowledges. This favours the view that it is a robust alternative, not a mere flicker, in Fischer's terms. And it certainly could be taken into account in an assessment and explanation of Joe's MR for his decision: the fact that Joe did not pay enough attention to moral reasons can worsen our moral assessment of his, in that it presents him as egoist and inconsiderate; and it can be made to weigh, at least partially, on explaining why, and to which degree, he is morally responsible for his decision. The alternative is not as such explanatorily irrelevant concerning Joe's MR. There is then reason to consider it as robust.

Pereboom (2009, p. 114) accepts, following a suggestion of mine (Moya 2006, pp. 65–66), that alternatives such as the one Joe has may have some weight in assessments and explanations of an agent's MR. They can improve or worsen, depending on particular circumstances, our moral evaluation of her and her acts. However, according to him, this is not sufficient for an alternative to be robust. He strengthens considerably Fischer's conception of the robustness of an alternative. A robust alternative has to be under the agent's control and be relevant to explaining her MR, as Fischer says. But, in order to be thus relevant, Pereboom contends, the agent has to understand (or at least have some degree of cognitive sensitivity to the fact) that, by choosing it, she would be, or at least would likely be, precluded from

the MR she now bears for what she decided and did (Pereboom 2009, p. 112; cf. 2001, p. 26, and 2003, p. 194).⁵ For these reasons, in Moya 2006 I dubbed robust alternatives in Pereboom's sense "exempting alternatives".

But why should we accept this strong concept of robustness? According to Pereboom, the intuition that lies behind the requirement of alternative possibilities for MR is "the 'off the hook' intuition: to be blameworthy for an action, the agent must have been able to do something that would have precluded this blameworthiness" (Pereboom 2009, p. 114).⁶ If, for example, we consider someone morally responsible for lying to us, we do this on the assumption that she could have not lied, so that, if she had not lied, she would not have been blameworthy. I have some worries about this thesis. Though the assumption of exempting alternatives may underlie many cases of ascription of MR, there are other cases in which we also take into account weaker alternatives, which would have assuaged, but not fully precluded, an agent's MR. We refer to these alternatives, which the agent could have chosen but did not, in order to explain, not why she is morally responsible in the first place, but why she bears a certain *degree* of MR. Our interest in this sort of robust, explanatorily relevant, though not exempting, alternative seems to cohere well with our view of MR as a gradual, and not just an all-or-nothing, property of human agents. This is also part of our intuitions about alternatives and MR. So there seems to be no principled reason to conceive of all robust, explanatorily relevant alternatives as exempting in Pereboom's sense. If someone harmed other people intentionally, an exempting alternative would be not to harm them; but learning that she did not care about those people's sufferings, or that she even scoffed at them, are not morally irrelevant pieces of information: they can be justifiably taken to aggravate the agent's blameworthiness. Thus this explanatory relevance of non exempting alternatives raises doubts about Pereboom's notion of robustness.

Let us, however, accept, for the sake of the argument, that the alternatives that should be available to agents in Frankfurt cases in order to save PAP are exempting alternatives. Now, going back to

⁵ Pereboom has revised progressively his characterization of robustness. I think the preceding paraphrase is faithful to his present conception of it.

⁶ Haji (personal communication) has doubts about this "off-the-hook" justification for PAP. He sees PAP as a control condition for MR. The idea is that responsibility requires plural control; if we did not have this sort of control, we could not make a difference to how our lives unfold.

Tax Evasion (2), even if Joe's attaining a certain level of attentiveness to moral reasons against evading taxes is relevant to explaining Joe's degree of MR for his decision, it does not seem to count as an exempting alternative. It is true that, had Joe freely attained the required level of attentiveness, the device would have been activated and Joe would not have been morally responsible for deciding to evade taxes. But the epistemic requirement for an exempting alternative would not have been met, for Joe, fully ignorant of the device, could not be reasonably expected to believe or understand that, just by attending to moral reasons against evading taxes, he would be exempted from his MR for deciding to evade them (and for doing so). The only alternative that Joe could reasonably believe that would allow him to get off the moral hook is just to refrain from deciding to evade taxes (and to act accordingly). But, of course, this exempting alternative was not available to him.

But let us pause a bit. As we have argued, becoming attentive to moral reasons is a morally relevant alternative. It is not like, say, catching a cold. Now, the reasoning that leads to the conclusion that it is not exempting seems to start from the assumption that, since Joe is fully unaware of the device in his brain and since it never gets activated, Joe's situation in *Tax Evasion* can be harmlessly assimilated, for the purposes of assessing his MR, to a normal situation in which there is no neurosurgeon and no lurking device, and in which Joe could have decided not to evade taxes. Now, if we judge Joe's case from this perspective, Pereboom is certainly right: attending to moral reasons is not an exempting alternative. For suppose that, after paying the required attention to moral reasons against evading taxes, Joe dismisses them and decides to evade taxes. It would be crazy to hold that Joe is not morally responsible for this decision because he attended to reasons against it. And Joe himself could not expect to get rid of blame by appealing to this mental act. Our intuitions, then, are clear in this respect.

However, the assumption that, concerning assessments of MR, Joe's situation can be assimilated to a normal situation, where no lurking device exists, is highly problematic: Joe's situation in *Tax Evasion* is not normal, for, even if the device never gets activated, its presence ensures that there are things that Joe *cannot* do. And our intuitions about when an alternative is exempting are highly sensitive to modal facts, to what an agent can and cannot do in particular contexts. As we have said, Frankfurt-inspired theories are "actual sequence" theories of MR, and so they tend to dismiss or devalue

modal facts; but these facts are very important to our pre-theoretic judgements about MR. Let me justify this with an example.

Suppose that someone is walking along a street and she suddenly witnesses an accident: a pedestrian is run over by a car and lies on the ground, with quite serious injuries; suppose further that the car driver absconds and that she is the only person who has witnessed the accident; she has the moral duty to help the victim; as it happens, she is a doctor, with a long experience in treating traumas, and has got a first aid case; what she does, however, is to take her mobile, dial an emergency number and ask for an ambulance; she could additionally have examined the injured person in order to determine his condition and see how she could start helping him with his injuries before the ambulance arrives, but she just feels tired and not in the mood to do that.

I think our judgement about this example is that the doctor bears some degree of blame because there is something more she could have done in order to help the victim and did not. This something more was an exempting alternative; had she chosen to do it, she would not have been blameworthy, and she understood that she would not; what she did, though better than doing nothing, is not enough to exempt her from blame; she ought, and could, have done better. But think of the following counterfactual variation of the story: things happen as in the original example, but now the witness is not a doctor, but a lay person, with no medical knowledge or training at all. In this counterfactual story, the witness of the accident would not have been blameworthy; the alternative she chose (to dial an emergency number and ask for an ambulance) exempted her from blame, for there was nothing she could additionally do to help the victim.

If we share these judgements, we can see that the same way of behaving exempts the agent from blame in the counterfactual story, but not in the original story. In the latter case, the agent had an exempting alternative (and knew she had it) which she could have gone for, but did not.

It seems, then, that the question whether, in a situation of a certain kind, a particular way of acting is an exempting alternative cannot be correctly answered without taking into account (among other things) what the agent can and cannot actually do in the circumstances. Raising the level of generality, my suggestion is that our judgments about these questions are guided (among other things) by the following principle:

(C) If someone cannot reasonably do more than she actually does in order to fulfil her moral duties, she is not morally obliged to do more, and so she is not morally blameworthy for not doing more.

In fact, (C) is formed by two conditionals. The first (“If someone cannot... she is not morally obliged...”) is roughly the contraposition of an old moral principle, namely that “ought implies can” (OIC). The second (“If someone is not morally obliged... she is not morally blameworthy...”) states, plausibly enough, that moral blameworthiness for A-ing (not A-ing) requires moral obligation not to A (to A). The implicit application of (C) seems to explain our judgements about the preceding example, both in the original and the counterfactual version.⁷

In order to deepen our enquiry, and before coming back to Pereboom’s *Tax Evasion*, we should take into account the subjective cognitive state of agents, for sometimes what we think we can (or cannot) do and what we actually can (or cannot) do are not coextensive. This is the case with Pereboom’s *Tax Evasion*, for Joe believes he can decide not to evade taxes, but he is wrong about this. And this will also be the case with Widerker’s *Brain-Malfunction-W*, which we will examine below.

3. *Awareness and Ignorance*

Let us go back to the original version of our example. Suppose that, after the doctor omits giving the injured man her personal medical help and he is already within the ambulance, she discovers with surprise that the case she has got is not her first aid case, but a similar but useless case that she has confused with it. She might then claim⁸ that she was not to blame for not giving first aid to the victim, for she could not have given it to him. Of course, given her ignorance about the content of her case, this does not preclude her blameworthiness, but it has the effect of lowering the standards for an alternative to count as exempting. By OIC, given that she could not have personally

⁷ Note that it does not follow logically from (C) that if someone can reasonably do more than she does, then she is ipso facto morally obliged to do it (though it may plausibly raise an expectation that she is). I think this is a positive trait of (C), which otherwise would burden us with lots of moral duties we would be unable to discharge, for in many cases we can do more, even if we do much. The qualification “reasonably” is important, anyway, as it is attending to the features of particular cases.

⁸ Implicitly applying both PAP and OIC.

aided the victim, aiding him was not morally required of her any more; but this does not preclude her from MR, for she ought and could have *tried* to aid him, which implied at least opening her case, if only to discover that it did not contain any medical material. In this situation, honestly trying to help the victim by opening the case would have been an exempting alternative. To see the mechanism implicitly at work here, imagine now that, in addition to containing useless stuff, the case's lock was actually stuck and she could not have opened it. Again, this modal fact lowers further the standards for exempting alternatives in the circumstances. It would have been enough, in these particular circumstances, for the doctor to get rid of moral blame, that she had (honestly though unsuccessfully) tried to open her case, which implied making certain (rather obvious) physical efforts. According to (C), in these circumstances, she could not reasonably have done more than this in order to fulfil her moral duties, so that, if she had done it, she would not have been morally obliged to do more and would have been precluded from blame. In the circumstances, honestly trying to open the case would have been an exempting alternative. If we assume that these were the circumstances in the example, it is plausible to hold that the doctor was not totally blameless, either, for she did not even try to open her case.

A result of these considerations is that, in cases of ignorance of inability, our judgements about the exempting character of an alternative rest on a "next best action" basis, in the following sense:

(NBA-ign) If, unbeknownst to her, an agent cannot do something A such that, if she did it, she would fulfil her duty and would be precluded from blame (and she knows that she would), then, in order to be so precluded, she should perform the next best action that reasonably was in her power to perform in order to fulfil her duty, where "the next best action" may be characterized, in general terms, as trying or attempting to A.

Which particular actions trying to A amounts to depends on the context, as we have seen in the example.⁹

⁹ "Trying to A" should not be understood as a purely mental act, in O'Shaughnessy's sense (O'Shaughnessy 1980); it should be taken to refer, in accordance with everyday usage, to ordinary ways of acting directed at A-ing; however, in particularly sophisticated contexts, which include some Frankfurt cases, it might refer to a mental act.

Dropping the assumption of ignorance of inability has some special consequences of its own, though the “next best action” basis holds here as well. Suppose, in effect, that the doctor definitely *knew* in advance that her case’s lock was stuck, because, shortly before the accident, she had tried in vain to open it. In this case, the standards for exempting alternatives are again lowered. We do not even require of her that she tried to open the case in order to preclude her from blame. In these circumstances, what she actually did in the example, namely to call for an ambulance, would probably be an exempting alternative, if this was the only thing she could reasonably do to fulfil her duty of helping the victim. With the assumption of knowledge of inability, NBA may be formulated thus:

(NBA-kn) If an agent knows (or justifiably believes) that she cannot do something A such that, if she did it, she would fulfil her moral duty and be precluded from blame, then, in order to be so precluded, she should perform the next best action that reasonably is in her power to perform in order to fulfil her moral duty.

Again, which action is the next best one is highly dependent on the context, but now we cannot characterize it as trying to A, for trying to A requires the belief that A-ing is not beyond one’s reach, which is not the case under the assumption of knowledge of (or justified belief in) inability. (NBA-ign) and (NBA-kn) look like corollaries or plausible extensions of principle (C).

If the preceding considerations are on the right track, they should have important consequences for which exempting alternatives exist in Frankfurt cases, since in these cases what the agent can do is severely restricted. Let us then go back to Pereboom’s *Tax Evasion* and apply the foregoing criteria to it. This is clearly a case of ignorance of inability. Joe believes that he can decide not to evade taxes, but this belief is false. Though he cannot make it, this decision is such that, if he made it, he would fulfil his duty and would be precluded from blame, and he knows that he would. So, by (NBA-ign), he should perform the next best action that reasonably was in his power to perform in order to fulfil his duty. According to (NBA-ign) “the next best action” may be characterized, in general terms, as trying to A. Now, what could “trying to decide not to evade taxes” amount to in this context? It is not easy to answer this question, but it looks plausible to say that part of the answer is: to gather evidence and reasons in favour of a decision of this kind and to pay

due attention to them. This is something that, as Pereboom himself acknowledges, Joe could voluntarily have done. But now we can see that, against his contention, it can count as an exempting alternative. According to (C), and provided that this was everything that Joe could reasonably do, in the circumstances, in order to fulfil his moral duty, he is plausibly taken to be morally blameworthy for his decision to evade taxes partly because he did not do the above. So, Joe is morally responsible for his decision to evade taxes and for evading them partly because he did not do everything that was reasonably in his power to do in order to honour his moral obligations: he ought to, and could, have thought of, and paid more attention to, moral reasons against deciding and acting as he did in order not to decide and act that way, but did not. He showed disrespect for morality, which he could have respected. And this is partly why he is morally blameworthy. In normal circumstances, with no device lurking, the standards for exempting alternatives would have risen to deciding not to evade taxes and not evading them; merely attending to moral reasons would not have been enough; but, since Joe could not have decided and acted that way, the standards lower to the next best action he could perform in order to fulfil his moral duties, which so becomes an exempting alternative.

So, on this plausible interpretation of the notion of an exempting alternative, *Tax Evasion* and structurally similar examples do not refute PAP: the agent is morally blameworthy, but, against appearances, he has robust, even exempting alternatives after all.

If the preceding considerations are on the right track, Pereboom's epistemic requirement on exempting alternatives looks too demanding and is in need of some reform to cover cases of ignorance of inability. In situations where an agent is unaware of her inability to perform an action that she correctly thinks would exempt her from blame, a "next best action" can be for her an exempting alternative, even if, not knowing that a better action is impossible for her, she does not believe that simply performing that next best action will make her blameless. This is what happens with our doctor when she is unaware that the case she is taking with her is not a first aid case: she does not believe that simply opening the case would exempt her from blame, but it would nonetheless. And this is also the case with Joe in *Tax Evasion*: he does not believe that simply becoming attentive to moral reasons against evading taxes would exempt him from blame, but it would, and with good reason, if our considerations are correct, for this is everything he could reasonably have done, in the context he was in, in order to fulfil his moral duty not to (decide to)

evade taxes. This is what trying to decide not to evade taxes would actually amount to in these circumstances and what Joe should, and could, have done.

4. *Pereboom's New Versions of Tax Evasion*

In response to criticisms, Pereboom has designed (at least) two new versions of his example. Let us focus on the first (2009, p. 117), though I will add to it some details he includes in the second (forthcoming) in order to strengthen its dialectical structure. On this new version, Joe is “aware and sensitive to the moral reasons not to evade taxes”; however, in the circumstances he is in, these moral reasons are overridden by self-interest. In fact, in such circumstances, and unbeknown to Joe, for him to decide not to evade taxes it is causally necessary “that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so [i.e. for evading taxes], a mental state he can produce voluntarily” (2009, p. 117). However, if the amount to be evaded were substantially higher, then he would decide not to evade taxes for moral reasons alone, even without imagining being punished. As in previous versions, this imagining, though causally necessary, is not causally sufficient for Joe to choose not to evade taxes. However, to ensure this choice, if the device in his brain were to sense the imagining, “it would electronically stimulate the right neural centers” so that Joe would inevitably choose to evade taxes. Again, “Joe does not imagine in this way being punished, and he decides to take the illegal deduction while the device remains idle” (2009, p. 117).

There are some differences between this version and the preceding one. In the latter, *Tax Evasion (2)*, the causally necessary condition for Joe’s deciding not to evade taxes, namely paying more attention to moral reasons, was clearly something morally better that Joe could reasonably have done in order to fulfil his moral duties; and, even if Joe was ignorant of the causally necessary character of that condition, it was reasonable to hold that he ought to have met it, as it was a natural step towards the (morally right) decision. In the new version, however, the corresponding causally necessary condition, namely to imagine being severely punished, is not clearly something morally better that Joe could reasonably have done and, given that he was ignorant that this was causally necessary for his making the right moral decision, it is not reasonable to hold that he ought to have imagined the punishment. This difference, however, is not enough for this new version to circumvent

principles such as (C) and (NBA-ign). Since Joe was aware of, and sensitive to, moral reasons against evading taxes, why should we accept that Joe could not have made the effort to pay more attention to those reasons and to decide according to them? Given that, in this situation, for him to decide against evading taxes it was causally necessary that he imagined being punished, this effort would have been powerless unless accompanied by that imagining, but the effort is something more which he could have done to fulfil his moral duties. There was, then, open to Joe a “next best action” after all, which he didn’t perform but could have. And, in accordance with (C) and (NBA-ign), it was an exempting alternative.

A second objection is this. Pereboom depicts Joe (or Joe’s abilities for practical reasoning) as reasons-responsive, in the sense that he can respond to stronger moral reasons than those he actually considers even without imagining being punished. That Joe has this ability is important in order to avoid suspicions about the soundness of his capacity for practical and moral reasoning, which could in turn raise doubts about his MR. But this feature of Joe’s has a cost. For if he is able to decide not to evade taxes for stronger moral reasons (if, for example, the amount to evade were higher) with no need of the punishment thought, it is then very hard to accept that it is literally *causally impossible* for him to make that morally right decision only for the moral reasons he actually considers, without such a thought of punishment. (And then he could have decided not to evade taxes, for the sign for the device’s firing, namely the imagining, would have been absent.) It is one thing to say that, given his self-interested character, it is very hard for Joe to give moral reasons pre-eminence over self-interest and very unlikely that he would do so. It is another thing to say that it is causally impossible for him to do this. It is very frequent that deciding in accord with moral reasons requires a greater effort of will than doing it according to self-interest. But this does not mean that making this effort is beyond the agent’s powers, if he is morally reasons-responsive. Under this assumption, the stipulation that it is causally impossible for Joe to decide not to evade taxes for moral reasons alone does not look realistic. Why should we accept this if Joe is sensitive and responsive to moral reasons? He should have decided against the tax evasion on the sole basis of the moral reasons he was aware of, with no need of the punishment thought, and there is no clear reason to think that it was causally impossible for him to do so. Stipulating that it was is not enough to make the

example psychologically convincing and credible enough to succeed against PAP.

However, let us accept, for the sake of the argument, that the punishment thought is actually causally necessary for Joe's deciding against evading taxes, so that without this thought it is causally impossible for him to make that decision. Now, since Joe is ignorant of this fact, it is not reasonable to hold that he ought to have imagined the punishment. Joe has no reason to suspect that the imagining is causally necessary for him to decide against the tax evasion. But then it seems that the possibility of Joe's making this right moral decision depends on a fortuitous event, namely the possibility of his vividly imagining being severely punished. That decision, then, is not appropriately under Joe's rational-cum-causal control. It is strongly dependent on luck. And this raises serious doubts about Joe's blameworthiness for not making it.

For these reasons, I think that this new version of *Tax Evasion* is also powerless to refute PAP.

5. *Widerker's Post-Dilemma Example*

This is Widerker's Frankfurt-style, post-dilemma example:

(*Brain-Malfunction-W*) Jones is deliberating as to whether to keep the promise he made to his uncle to visit him in the hospital shortly before his uncle is about to undergo a critical operation. Jones is his uncle's only relative, and the visit is very important to the uncle. The reason for Jones's deliberating is that, on his way to the hospital, he (incidentally) met Mary — a woman with whom he was romantically involved in his distant past, and whom he has not seen since then. Mary, being eager to talk to Jones, invites him for a cup of coffee in a nearby restaurant. She explains that she is in town just for a couple of hours, and wishes to spend those hours with him. Jones is aware that if he accepts Mary's offer, he will not be able to make it to the hospital during visiting hours. Normally, one can avoid deciding as one does by deciding otherwise. But in our scenario Jones does not have that option, since shortly after beginning to deliberate, he undergoes a neurological change as a result of which one of the (neurological) causally necessary conditions for his deciding otherwise, a condition which we may call "*N*", does not obtain. It is also assumed that this fact is unknown to Jones (who believes that he can decide to keep the promise),

and that N 's absence does not affect his deliberation process. In the end, Jones decides on his own not to keep the promise, and spends the afternoon with Mary. (Widerker 2009, pp. 89–90; cf. Widerker 2006, p. 170)

In my 2007 paper I criticized this example on the following grounds: given that, shortly after Jones starts deliberating, N , a neurological necessary condition for Jones to decide to keep his promise, ceases to obtain, the apparatus of practical reasoning and decision making with which he faces his choice between staying with Mary and visiting his uncle is defective, in that it is not reasons-responsive: no matter how strong the reasons Jones might be faced with for deciding to visit his uncle, he still would not make that decision, for a neurological necessary condition for making it would be absent. And if, as seems plausible and is widely accepted, reasons-responsiveness is a necessary condition of MR, Jones should not be judged morally responsible for his decision. But, since a valid counterexample to PAP must feature an agent who (1) is morally responsible for his decision and (2) lacks robust alternatives to it, Widerker's example is not a valid counterexample in that it does not meet the first requirement.

I still think this criticism is correct.¹⁰ But, according to the main theme of this paper, I will try to add critical pressure on Widerker's example from the perspective of the second requirement, namely the absence of robust alternatives. I will try to show that Jones does have robust alternatives after all. I will attempt to do this partly on the basis of an interesting answer of Widerker's (2009) to my "no-reasons-responsiveness" criticism. None the less, the point might be made independently of this answer.

Widerker's response to my criticism starts from drawing a rather sharp distinction between deliberation and practical reasoning, on the one hand, and decision making on the basis of reasons, on the other (cf. Widerker 2009, pp. 92–93). Widerker accepts that reasons-responsiveness, or, as he puts it, the ability "to respond differentially to reasons", is a requirement of MR, but he thinks that this requirement can be met on the basis of a sound capacity for deliberation and

¹⁰ It may be contentious, however, whether N , or its lack, is part of the mechanism of deliberation and decision making with which Jones faces his choice. Manuel Vargas and Ishtiyaque Haji, independently, called my attention to this difficulty. However, ruling out alternatives by tinkering with *actual* brain properties, instead of *counterfactual* factors, is not a good idea anyway, for it raises doubts about the integrity of the agent's rational abilities, and so about his MR.

practical reasoning; there is no additional need for a faultless capacity for decision making. Part of what the former capacity amounts to is an agent's ability to form correct judgements or beliefs about what she would decide and do if she were to have certain reasons that she does not presently have. Widerker contends that Jones retains this ability, or at least that there are no good reasons to think that he does not; so, if we asked him what he would decide if there were much stronger reasons for keeping his promise, then he would answer that, in that case, he would decide to keep his promise and visit his uncle, rather than to stay with Mary (cf. Widerker 2009, p. 92). Even if it is true that, owing to the absence of N , Jones would not be able to make that decision, the fact that he can form beliefs that respond differentially to (weaker and stronger) moral reasons shows that he retains reasons-responsiveness, at least on the proposed construal of this notion.

This is an ingenious move. It is dubious, however, whether an agent who has a sound capacity for forming correct beliefs about what to decide and do given certain reasons but who is causally unable to translate these beliefs into appropriate decisions can count as being normal and competent enough, from the perspective of practical reason, not to raise doubts about her MR for her decisions. Normal, competent moral agents are usually able, barring sporadic episodes of weakness of the will, to make decisions that accord with their practical judgements or beliefs about what they have best or better reasons to do. However, Jones does not have this ability when he faces his choice. In other words, I think that the criticism based on reasons-responsiveness retains a lot of its force, Widerker's response notwithstanding.

However, as I announced, I do not want to pursue this line of argument further. Even if we accept Widerker's response and agree that Jones is morally responsible for his decision, I think that this response leads him to violate the second requirement for a successful Frankfurt case, namely the agent's lack of any robust alternative. Let me argue for this contention.

If we accept, with Widerker, that, in spite of N 's absence, Jones retains a sound capacity for forming correct beliefs and judgements about what he should and would do given certain reasons, and if we take him (as we should if we are to consider him as a moral agent) to be sensitive to moral reasons, then we are entitled to expect him to have formed the *practical judgement* that keeping his promise and visiting his uncle was the decision to take and the thing to do then

and there, in that it was better, from a moral point of view, than staying with Mary. We may legitimately assume that his actual decision in favour of staying with Mary reflected instead the opposed, and maybe implicit, practical judgement that this alternative was all-out preferable to visiting his uncle.¹¹ So, his actual practical judgement went against the demands of morality, and Widerker gives us no good reason for thinking that Jones could not have formed the morally correct judgement that visiting his uncle was the thing to do in the circumstances. It is true that, given the absence of *N*, Jones could not have made a decision that fitted that practical judgement but, in the circumstances of the case, forming that practical judgement was everything he could reasonably do in order to make that decision and fulfil his moral duties; so, by (C), if he had honestly tried to make the morally right decision, by forming that judgement, he would not have been morally obliged to do more, namely to decide to keep his promise, and would not have been morally responsible for not doing it. By (NBA-ign), provided that, unbeknown to him, making the decision to keep his promise was causally impossible for Jones, he should have formed the indicated practical judgement, for this was the next best action Jones could have performed in the circumstances; it was his (best) way of trying to make that decision. Of course, if *N* had not been absent, and Jones could have made the right moral decision, then forming the aforementioned practical judgment would not have been enough for precluding his MR; only this decision (and, let me add, his consequent action) would have been enough for precluding this; as things actually were, however, the practical judgement would have sufficed; had he honestly tried to decide to keep his promise by (sincerely) forming this practical judgment on the basis of his moral reasons, which he was sensitive to, he would have been exempted from the MR he now bears. So, on the assumption that, in spite of *N*'s absence, Jones retained a sound capacity for deliberation and practical reasoning, we should accept that he had an exempting alternative in Widerker's example.

A possible move for Widerker may be to modify his example in such a way that neurological condition *N* is causally necessary, not only for Jones's decision to keep his promise, as in the original version, but also for him to form the practical judgement that keeping

¹¹ One important reason for this presumption is that, if Jones had formed the opposite practical judgement, namely that visiting his uncle was the all-out preferable option, this would have brought in the example an inconsistency between his practical judgement and his actual decision, and so an element of irrationality that would have undermined intuitions in favour of his MR.

his promise and visiting his uncle was the decision to make and the thing to do in the circumstances. One problem with this move is that it threatens Widerker's response to the "no-reasons-responsiveness" objection, for, in this new version, N 's absence would negatively affect not only Jones's capacity for decision making, but also his capacity for deliberation and practical reasoning; on losing N , Jones would be causally unable to form the indicated practical judgement no matter how strong the reasons he might have or be presented with for forming it. Widerker might insist that Jones could still form correct beliefs about the right practical judgements he would form given certain stronger reasons for visiting his uncle. But even if this response were to succeed, a PAP defender might, on the basis of (NBA-ign), point out that Jones should have paid more attention and assigned more weight to the moral reasons he actually had for keeping his promise and visiting his uncle. This would bring Widerker's example closer to Pereboom's *Tax Evasion* (2), at least with respect to the alternative open to the agent. Widerker might then modify the example again, by making N also necessary for Jones to pay that degree of attention and assign that degree of weight to his actual moral reasons for keeping his promise. However, I think this dialectic cannot go on forever, for in the limit, if there is absolutely *nothing* Jones can do to fulfil his moral duties, if everything he can do to fulfil them is just nothing, then, by (C) (and OIC), he is not morally obliged to fulfil them any more and so is not morally responsible for not fulfilling them. PAP would remain safe at the end.

6. Conclusion

If the preceding considerations are correct, then neither Pereboom's nor Widerker's recent Frankfurt examples succeed as counterexamples to PAP, for in them the agent has robust, even exempting alternatives at his disposal. (As we have argued, there may be also problems related to reasons-responsiveness.) It is worth noting that principle (C) and its corollaries (NBA-ign) and (NBA-kn) are independent of a prior commitment to the truth of PAP; as we have seen, they are motivated by intuitively correct and eminently fair judgements raised by examples which reflect our implicit beliefs about the connections between ability, alternatives and MR.¹²

¹²This paper is part of the research project "Alternatives, Belief, and Action" (FFI2009-09686), supported by the Spanish Ministry of Science and Innovation, in collaboration with the European Regional Development Fund of the European Community. Several people have helped me to improve previous versions of this

REFERENCES

- Davidson, D., 1973, "Freedom to Act", in T. Honderich (ed.), *Essays on Freedom of Action*, Routledge and Kegan Paul, London, pp. 139–156. (Reprinted in D. Davidson, 1982, *Essays on Actions and Events*, Clarendon Press, Oxford.)
- Fischer, J.M., 2010, "Frankfurt Cases: The Moral of the Stories", *Philosophical Review*, vol. 119, pp. 315–336.
- , 1999, "Recent Work on MR", *Ethics*, vol. 110, pp. 93–139.
- , 1994, *The Metaphysics of Free Will*, Blackwell, Oxford.
- Frankfurt, H.G., 1969, "Alternate Possibilities and Moral Responsibility", *Journal of Philosophy*, vol. 66, pp. 829–839. (Reprinted in (and quoted from) *The Importance of What We Care About*, Cambridge University Press, Cambridge, 1988, pp. 1–10.)
- Ginet, C., 1996, "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing", *Philosophical Perspectives*, vol. 10, pp. 403–417.
- Haji, I., 2010, *Incompatibilism's Allure*, Broadview Press, Toronto.
- Haji, I. and M. McKenna, 2006, "Defending Frankfurt's Argument in Deterministic Contexts: A Reply to Palmer", *Journal of Philosophy*, vol. 103, pp. 363–372.
- , 2004, "Dialectical Delicacies in the Debate about Freedom and Alternative Possibilities", *Journal of Philosophy*, vol. 101, pp. 299–314.
- Kane, R., 1985, *Free Will and Values*, State University of New York Press, Albany.
- McKenna, M., 1997, "Alternative Possibilities and the Failure of the Counterexample Strategy", *Journal of Social Philosophy*, vol. 28, pp. 71–85.
- Mele, A. and D. Robb, 2003, "Bbs, Magnets and Seesaws: The Metaphysics of Frankfurt-Style Cases", in Widerker and McKenna 2003, pp. 127–137.
- Moya, C.J., 2010, "Blame, Moral Obligation and Alternative Possibilities", ms. presented as an invited lecture to the VI Conference of the Spanish Society for Analytic Philosophy (Tenerife, Spain).
- , 2007, "Moral Responsibility Without Alternative Possibilities?", *Journal of Philosophy*, vol. 104, pp. 475–486.
- , 2006, *Moral Responsibility. The Ways of Scepticism*, Routledge, Abingdon/New York.
- Naylor, M.B., 1984, "Frankfurt on the Principle of Alternate Possibilities", *Philosophical Studies*, vol. 46, pp. 249–258.
- O'Shaughnessy, B., 1980, *The Will: A Dual Aspect Theory*, Cambridge University Press, Cambridge, 2 vols.

paper: Carlos Patarroyo, Tobies Grimaltos, Derk Pereboom, and Sergi Rosell, among others. Very special thanks are due to Ishtiyaque Haji for long and detailed written remarks. I am also very grateful to three anonymous referees for *Crítica* for their comments, suggestions and criticisms.

- Otsuka, M., 1998, "Incompatibilism and the Avoidability of Blame", *Ethics*, vol. 108, pp. 685–701.
- Pereboom, D., 2009, "Further Thoughts about a Frankfurt-Style Argument", *Philosophical Explorations*, vol. 12, pp. 109–118.
- , 2003, "Source Incompatibilism and Alternative Possibilities", in Widerker and McKenna 2003, pp. 185–199.
- , 2001, *Living Without Free Will*, Cambridge University Press, Cambridge/New York.
- , forthcoming, "Optimistic Skepticism about Free Will", in P. Russell and O. Deery (eds.), *The Philosophy of Free Will: Selected Contemporary Readings*, Oxford University Press, New York.
- Widerker, D., 2009, "A Defense of Frankfurt-Friendly Libertarianism", *Philosophical Explorations*, vol. 12, pp. 87–108.
- , 2006, "Libertarianism and the Philosophical Significance of Frankfurt Scenarios", *Journal of Philosophy*, vol. 103, pp. 169–187.
- , 1995, "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities", *Philosophical Review*, vol. 104, pp. 247–261.
- Widerker, D. and M. McKenna (eds.), 2003, *Moral Responsibility and Alternative Possibilities*, Ashgate, Aldershot.
- Wyma, K.D., 1997, "Moral Responsibility and Leeway for Action", *American Philosophical Quarterly*, vol. 34, pp. 57–70.

Received: April 16, 2010; revised: March 8, 2011; accepted: March 9, 2011.