

Three Recent Frankfurt Cases

Robert Lockie

Received: 15 August 2013 / Revised: 19 November 2013 / Accepted: 11 April 2014
© Springer Science+Business Media Dordrecht 2014

Abstract Three recent ‘state of the art’ Frankfurt cases are responded to: Widerker’s Brain-Malfunction-*W* case and Pereboom’s Tax Evasion cases (2 & 3). These cases are intended by their authors to resurrect the neo-Frankfurt project of overturning the Principle of Alternative Possibilities (PAP) in the teeth of the widespread acceptance of some combination of the WKG (Widerker-Kane-Ginet) dilemma, the Flicker of Freedom strategy and the revised PAP response (‘Principle of Alternative Blame’, ‘Principle of Alternative Expectations’). The three neo-Frankfurt cases of Pereboom and Widerker shown to be insufficient for their intended purpose. Of central importance to any account of responsibility is that this applies at the level of the Right and not the Good. Arguments of Carlos Moya are expanded and augmented by considerations from Chisholm, Lucas, Dummett and Lockie (2003) to show that a number of severe problems remain for anyone attempting to resurrect the Frankfurt project.

Keywords Moral responsibility · Alternative possibilities · PAP · Frankfurt · Flicker of freedom · Right versus Good

A vast Frankfurt literature has developed over the last 45 years, and at the time of writing the leading pro-Frankfurt arguments are Widerker’s *Brain-Malfunction-W* case and the later refinements out of a series of versions of Pereboom’s *Tax Evasion* case (versions 2 and 3). This paper responds to these three cases. It argues that they miss their mark and that the Frankfurt literature still fails to overturn the thesis that there is a fundamental leeway requirement on free will and / or moral responsibility¹. In so doing

¹“Free will and / or moral responsibility...” Henceforth I shall restrict myself to the latter, as this is where the Frankfurt debates are most centrally situated; but this is to flag that one may also be concerned with these debates as they apply to the former, and may be thus concerned not simply from the standpoint of an incompatibilist—the semi-compatibilist position precisely came to diverge from more mainstream versions of compatibilism in conceding that we could not be *free to do otherwise* in a determined world, yet could be *responsible*. Only in recent decades have these debates commonly circumvented the issue of metaphysical freedom to engage with issues of responsibility direct.

R. Lockie (✉)

Psychology Department, University of West London, Paragon House, Brentford TW8 9GA, UK
e-mail: bob.lockie@uwl.ac.uk

the paper draws some general morals as to the unlikelihood of the Frankfurt literature ever achieving its original aims: morals drawn from connected themes that emerge concerning the *right* vs the *good*, *deontology* vs *consequentialism*, the *intensionality of intentions* and the essential relevance of *epistemic awareness* to *responsibility properly-so-called*².

This first section summarises the state of play prior to these cases emerging. The second section introduces and makes a first response to Widerker's counterexample. The third (long) section introduces, disambiguates and responds to Pereboom's two counterexamples and assimilates these to Widerker's case. The fourth section makes a point (after Lucas 1970) against a crucial move deployed in all these examples. The fifth section flags a series of problems with accepting the terms of these examples and with issues of obscurity / ambiguity generally in this literature. The sixth and seventh sections deal exhaustively with issues (issues already flagged and employed in the rest of the piece) concerning the right and the good and the intensionality of intentions: in so doing, relating these highly specialised debates to issues central to moral theory and action theory more generally.

Prior to These Cases

The basic Frankfurt scenario involves a *Victim*, an *Agent* and an enormously powerful and knowledgeable neuroscientist, *Intervener*. Agent is potentially about to shoot Victim of her own accord. Intervener monitors Agent at a distance with great accuracy. Were Agent not to be about to decide to shoot Victim, Intervener would come to know and step in to effect the shooting himself—say, through direct brain stimulation to produce innervation of Agent's trigger-finger. This proves unnecessary as Agent decides to shoot Victim entirely without manipulation. We are to conclude that Agent is morally responsible for shooting Victim even though she could not have done otherwise—hence that 'could have done otherwise' is not necessary for moral responsibility: the Principle of Alternative Possibilities (PAP) is false.

The Frankfurt literature had been taken to have foundered—either to stalemate or defeat—substantially on three interrelated responses: the flicker of freedom response, what I shall call the 'intentional PAP' response and the WKG dilemma. The three recent cases we will shortly examine try to reanimate the anti-PAP literature in the face of these responses.

Flicker of Freedom

This response is that the freedom possessed by the putatively responsible Agent was that she could have been going to decide to refrain from killing Victim. Had she not

² Some of the core arguments of this paper are prefigured in the important work of Carlos Moya (2006, 2007, 2010, 2011)—which I commend to the reader and have fully acknowledged in the text—along with the work of Otsuka, Widerker-Kane-Ginet & Copp. However, these arguments are augmented by and integrated with original contributions involving material in sections below entitled "Necessary, Sufficient & Counterfactually Uncertain", "Self-Absorption and the Aims of Responsible Conduct", "The Intensionality of Intentions" and "Non-moral Reasons? Why We Won't be Considering Tax Evasion 3 (Much)"—and throughout, the distinction between the right and the good (especially the section entitled "Robustness, Intensionality, Right vs Good: A Tension"). However, before getting to these original contributions it will be necessary to rehearse and reincorporate Moya and others' important work.

been going to decide to kill Victim, by hypothesis Intervener would have known and acted (via innervations to Agent's musculature, for example). But then, the flicker response has it, this would have absolved Agent from moral responsibility. Agent is responsible because she did not begin to act (or begin to decide to act) in such a way that Intervener rather than Agent stepped in to be responsible for Victim's death—that was Agent's alternative possibility, her *flicker of freedom*. A standard response from the Frankfurters is that this state of affairs (the possibility of a precursor to decision on Agent's part: the 'flicker' which would have triggered Intervener's intervention were Agent to have started to initiate it) is *insufficiently robust* to ground Agent's moral responsibility. For instance, this flicker might just be a tiny physiological sign (Agent blushing red, for example).

Intentional PAP

Strictly, the Frankfurt cases do establish that PAP is false, but this is a superficial result, since a different way of couching the leeway requirement on responsibility (one that is more faithful to the deep motivations underlying that principle) remains true. As described, the Frankfurt case stipulates that there is no alternative to *Victim's death*—so, there is moral responsibility yet no alternative when this be described purely at the level of such a consequentialist outcome. Nevertheless, it remains true to say that Agent may be *held responsible* only if there was an alternative to *Victim's murder by Agent*. PAP is overturned only to the extent that Victim *dies* either way, but Agent doesn't *murder* either way. A way of expressing this that effects connections with moral theory more generally is that PAP may at best be overturned at the level of the good, but not (crucially) at the level of the right—of which more below. Intentionally, the difference between Victim murdering and Intervener murdering is a huge one—a difference which, at the level of action and agency, hence moral responsibility, is most certainly 'robust'. And this is true however behaviourally or neurologically inconsequential a flicker it may from Agent which constitutes the sign that would have led Intervener to act—and in turn, regardless of how causally inconsequential that flicker would be for Victim (who is dead either way). This anti-Frankfurt insight is then captured in a neighbourhood, intentional, leeway principle: Otsuka's (1998) 'Principle of Avoidable Blame' (PAB), Widerker's (2000) 'Principle of Alternative Expectations' (PAE), Moya's (2011) 'doing everything one can' (DEC) or Copp's (2003) Argument from Fairness.

The WKG Dilemma

Is Intervener held to be omniscient and omnipotent about a matter of future decision over the actions of the putatively responsible Agent? If he is, the question is begged against the Frankfurter's incompatibilist libertarian opponent. If, though, Intervener isn't omniscient and omnipotent, then the thought experiment doesn't get off the ground. This, telling, objection has come to be referred to as the 'WKG Dilemma': (Widerker 1995; Kane 1996; Ginet 1996).

These three responses represent, in concert, a powerful rebuttal of several generations of Frankfurt literature. The three Frankfurt cases we start to consider below aim, in the teeth of such responses, to resurrect the case that leeway (whether PAP proper or

one of the neighbouring principles) is not required for moral responsibility. Before getting to these cases however, it is worth making a brief note about the epistemic status of the principle that is here in dispute, and the consequent metaphilosophical complexion of the dialectic which surrounds this dispute.

The Metaphilosophical Framework and the Logic of the Dialectic

Frankfurt, in his original paper, was I think too modest in stating of his target that “It [PAP] has generally seemed so overwhelmingly plausible that some philosophers have even characterised it as an a priori truth” (Frankfurt 1969/1988: 1). The defender of PAP (or cognate principle) typically takes this principle to be a priori—or at least self-evident in some strong sense. One may of course claim that a putatively a priori (or at least strongly self-evident) principle is false, and Frankfurt’s thought experiment is intended to do just that: to reveal that PAP is not true hence, a fortiori, not a priori. To avoid begging the question, Frankfurt arguments attempt to reveal that PAP is not true by working from premises the libertarian has no grounds, internal to his position, to reject (so, for example, without assuming the truth of global determinism—which would be one horn of the WKG dilemma). The extended dialectic over Frankfurt arguments—for and against—has then mostly not involved anti-Frankfurters putting forward pro-PAP arguments (though there are such arguments in the literature). This dialectic has mostly involved anti-Frankfurters attempting to rebut the negative arguments of the Frankfurters—which negative arguments are seeking to advance from premises or presuppositions which the leeway incompatibilist’s starting position does not internally motivate him to dispute. Perhaps this framework of dialectic is unfair. Perhaps the anti-Frankfurt philosophers should have to positively advance arguments for their position. Perhaps even they should have to advance such arguments from premises their Frankfurt, compatibilist, opponents will also accept—after all, the Frankfurters seek to do that with them. If, though, they see (as I do) PAP (or a cognate principle) as a priori, one should be careful in requiring of them positive arguments. That is the whole point of a claimed a priori truth—that it is a priori. This is surely not (not yet, at least) for the anti-Frankfurt philosophers to beg the question against their opponents. A putatively a priori truth can be opposed by thought experiments (like Frankfurt’s) or perhaps other counter-arguments. But a *positive* argument for a claimed-as a priori truth might at best perhaps involve Bealer’s (1996a, b) ‘rephrasal strategies’—or some other attempt to articulate the a priori intuition in other ways—it precisely can’t involve constructive argument which goes much beyond this: that’s just what it is to be a (claimed) a priori truth.

Now there are arguments for PAP (or cognate principles) in the literature. For example, flagged already (in the section entitled “**Intentional PAP**”) were Copp’s (2003) Argument From Fairness and Otsuka’s (1998) PAB—and there are others³. Whether these are ‘rephrasal strategies’ or constructive arguments in their own right would be an issue that would take us too far from the topic of this paper: an attempt to rebut the most recent generation of anti-Frankfurt arguments, not an attempt (positively) to advance a defence of PAP (or cognate principle); which principle I claim, in line with Frankfurt’s

³ One that is rather different to those mentioned already is Copp’s (2003) Argument From Action-Guidingness—that the point to moral requirements is to guide agents’ decisions among their alternatives.

characterisation of his opponents, and without further (positive) argument, to be either a priori or so strongly self-evident as to require no further positive argument. If the Frankfurt theorist cannot successfully advance arguments working from within premises that the libertarian is already committed to, but the PAP theorist cannot either, then perhaps the proponent of PAP does not win his case, perhaps things rather end in Pereboom's (2009: 116, 2012: 306) "dialectically unsatisfying position"—a kind of metaphilosophical stalemate not without precedent in philosophy. However, it is indisputable that the Frankfurt tradition has in fact sought to refute PAP whilst advancing from premises that the libertarian has no grounds, internal to his position, to dispute. That was the whole, 45 year promise that this literature held out. It has involved the ongoing development of Frankfurt examples whose architects did not seek to dispute the leeway-libertarian framework presuppositions of their opponents; rather working (only for the sake of argument) within these presuppositions to seek to establish avowed moral responsibility in a (conceded) libertarian universe in which *in this instance* the ability to do otherwise could not reasonably be disputed—for reasons anchored in the details of the *recherché* thought experiment in question.

This paper argues that the most recent three state-of-the-art arguments from the Frankfurt tradition fail in this regard—and draws some more general morals therefrom. I am here attempting to rebut arguments, based on thought experiments, which see themselves as moving from premises and intuitions that the (principled) libertarian defender of PAP has no grounds, internal to that position for rejecting, to a conclusion that responsibility is present despite the fact that the agent could not have done otherwise; and which (were they to succeed) would a fortiori establish that neither PAP nor any cognate principle had the epistemic status its defenders claim (of a prioricity, or at least strong self-evidence).

Widerker's Brain-Malfunction-W

The first of the three recent Frankfurt cases we consider originated with David Widerker. Widerker is an incompatibilist libertarian, but one of several who now think that a version of Frankfurt's anti-PAP argument is sound; and thus that the libertarian should not locate his libertarianism in a defence of PAP. Widerker instead defends a non-PAP source incompatibilism (there are pro-leeway source incompatibilists also—the present author is one). Widerker's central case against PAP was first advanced in Widerker (2006) and is as follows.

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. ... 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: 89–90).

Widerker subsequently states

The term 'on his own' is to be understood in a libertarian sense, implying that such a decision is neither nomically determined nor caused (Widerker 2011: 269).

Decide vs Deliberate

Why can't we run a 'flicker of freedom' counter to Brain-Malfunction-W? Ultimately I shall endorse and extend an argument of Moya (2011) to claim that we can indeed do this; but Widerker takes this avenue to be closed to us. Widerker holds that Jones (a) cannot *decide* otherwise because of the neurological stymie $-N$, (so, no 'flicker'—or at least no 'robust' flicker); yet (b) this stymie does not affect Jones' *deliberation* process. Because of (b) Jones may be held *morally responsible*; yet because of (a) Jones *couldn't do otherwise*.

One point of note is that for Widerker, (a) 'decide' and (b) 'deliberate' have now become terms of art—or at least, placeholders for a committed distinction. 'Deliberate' is an earlier-stage, more-purely-mental thing, uncoupled (radically) from action; while 'decide' is more of a late-stage, actional (or on the way to actional) thing, something which occurs, unless stymied, after neurological stage N . The need for Widerker to draw this distinction became clear after Moya (2007) in a separate, non-Flicker-based counter-argument—one that will not concern us—tried to maintain that Jones could not be seen as morally blameworthy, if, as Widerker's example requires, the stymie on neurological state N meant that Jones was not capable of responding to reasons (reasons to visit his uncle rather than stay with Mary). To escape Moya's (2007) argument, Widerker (2009) then explicitly emphasised the distinction noted, between:

- a. Jones' capacity for *decision-making on the basis of reasons* (which Widerker agrees is impaired by neurological condition $-N$). In Widerker (2011) this has become 'to decide': n.b. *decide* is now a term of art.
- b. Jones' capacity for *practical reasoning* (which Widerker claims is not impaired by neurological condition $-N$, because "this ability need not be cashed out in terms of .. the different decisions the agent would be able to make .. in different circumstances" Widerker (2009: 93)). In Widerker (2011) this has become 'to deliberate': n.b. *deliberate* is now a term of art.

Let us suppose that this distinction is not hostage to empirical, neurophysiological, fortune, and proceed with it. In Brain-Malfunction-W a necessary condition (N) on Jones deciding to visit his uncle is removed (so, no alternative possibility) but a

sufficient condition for not-deciding-to-visit his uncle isn't put in place (so, moral responsibility is retained). Points: logically this is incoherent unless some substantial qualifications are made. Look at 1 & 2 below and compare to 3 & 4:

1. If a thing is a dolphin then it is a mammal [M].
2. If a necessary condition for this thing's being a dolphin is removed [$-M$] then a sufficient condition for this thing's not being a dolphin is put in place.
3. If Jones decides to visit his uncle, then a neurological condition obtains [N].
4. If a necessary condition for Jones deciding to visit his uncle is removed [$-N$] then a sufficient condition for Jones not deciding to visit his uncle is put in place.

So the qualification must be that a *causally* not logically necessary condition on Jones deciding to visit his uncle is removed (this removal being neurological stymie $-N$), but with a causally sufficient condition of Jones deciding to visit his uncle not put in place in turn. Except that this appears unsustainable too:

5. A causally necessary condition of catching a cold is infection with the cold virus [V].
6. If all cold viruses are removed [$-V$] that is indeed causally sufficient for my not catching a cold.

So we have to qualify things further, thus: a causally necessary condition (N) for Jones deciding to visit his uncle is removed, but a causally sufficient condition for Jones *not-starting-to-choose-to-decide-to-visit his uncle prior to stymie $-N$* (i.e. not *deliberating* [q.v.] to do this) isn't put in its place; only a causally sufficient condition for Jones *not-succeeding-in-choosing-to-decide-to-visit his uncle* is put in its place (i.e. *not deciding* [q.v.] to do this).

Jones isn't compelled to choose to stay with Mary ('choose' for me, will not be a term of art). He's on the one hand libertarian-free to go all the way to choosing to stay with Mary (he can *deliberate* and then *decide* this). Or, in his practical reasoning, he is libertarian-free to go all the way to deliberating to visit his uncle; but he can go only this, deliberative, part of the way to deciding to visit his uncle—only as far as stymie $-N$. This is as far towards choosing to visit his uncle as he can get. *But then a straightforward 'flicker' response is available*: Jones is morally responsible for not trying to do what's right, up to the point that he would have ended up butting his head against neurological stymie $-N$ (notwithstanding that he doesn't know that this, as an option, is all that's available to him, notwithstanding that intentionally, from within his moral-epistemic perspective, he would not merely have been aiming to snarl himself in a neurological stymie: cf. "The Intensionality of Intentions" section below).

Were Jones to start upon the path that could eventuate in the choice to visit his uncle he could do so under his own steam, as it were ("in a libertarian sense"). Enough of the precursor to that uncle-visiting-proto-action/decision would poke its head out (prior to stymie $-N$) for us to... what? Surely for us to judge Jones as then not morally responsible for failing to decide to visit his uncle—as he'd done all he could to go in that direction prior to stymie. Enough for a flicker of freedom, a *robust* flicker that is—one that cannot eventuate in action, and (by stipulation of the thought experiment) cannot even eventuate in *decision*; but can absolve him of moral responsibility.

Jones is and must be described as deliberating freely “in the libertarian sense” as to whether to stay with Mary or visit Uncle (to avoid one horn of the WKG dilemma). He’s described as *not* free (in any sense) to *decide* (an actionally-coupled term of art)⁴ to visit Uncle. But since he is free to deliberate⁵ to visit Uncle he’s morally responsible for not doing this. Deliberating to visit Uncle was his alternative possibility. To defend the example with the familiar response that this proto/precursor-to decision (‘deliberation’) is not *robust* enough as an alternative, is to say Widerker’s example doesn’t even get off the ground, for whatever ‘deliberation’ is, it is the component of Widerker’s thought experiment that is and must be robust enough to vindicate attributing sufficient moral responsibility to his agent, Jones: to blame him for not doing as he ought. (And as will be seen, there are general morals to be drawn here about the suspicious nature of the ‘robustness’ counter through the history of Frankfurt cases—*cf.* the section entitled “Robustness, Intensionality, Right vs Good: A Tension” below).

Credit for this ‘flicker’ objection to Brain-Malfunction-W is owed to Moya (2011). This example raises important further issues which will be considered after the initial encounter with Pereboom’s cases.

Pereboom’s Tax Evasion 2, Tax Evasion 3

Pereboom’s Tax Evasion 3: Joe is considering claiming a tax deduction for the registration fee that he paid when he bought a house. He understands that claiming this deduction is illegal, but he believes that he probably won’t be caught, and that if he were, he would plead ignorance and likely [sic] do so successfully. 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. He is aware and sensitive to the moral reasons not to evade taxes in the way he is contemplating, but in this case these reasons are overridden by self-interest. If he were to refrain from deciding to evade taxes, it would be for self-interested reasons. In fact, it is causally necessary for his failing to decide to evade taxes in this situation that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so, a mental state he can produce voluntarily. (He knows that his jurisdiction allows for a prison sentence for the sort of tax evasion he is contemplating.) However, Joe’s imagining in this way being punished is not causally sufficient for his failing to choose to evade taxes. If he were to do so, he could, exercising his libertarian free will, either decide to

⁴ Contrast the position that Widerker’s thought experiment leads him to, with the position of such figures as Owens (2000) and many others in the epistemic literature. Widerker’s actionally-coupled freedom of *decision* is seen by him as specifically problematic—in ways in which his more purely mental *deliberation* is not. Yet for a really major strand of thought within the epistemic literature this view of matters is precisely inverted: a relatively unproblematic freedom of action is specifically contrasted with a supposedly impossible aspiration to attain a freedom of thought. (*Cf.* The doxastic voluntarism debates, among others). This matters because the free will debates are starting to have a major impact in the epistemic deontology debates—debates in which issues of ‘ought’ implies ‘can’ are of central importance.

⁵ To say “I deliberate to Φ ” is an odd locution, (“deliberate” sounds like a process, not a success state) but it’s a locution forced on me by Widerker’s choice of vocabulary.

evade taxes or refrain from so deciding (without the intervener's device in place). However, to ensure that he will decide to evade taxes, a neuroscientist has, unbeknownst to him, implanted a device in his brain, which, were it to sense his imagining with a certain degree of vividness being at least fairly severely punished for evading taxes, would electronically stimulate the right neural centers so as to inevitably result in his making this choice. As it happens, Joe does not imagine in this way being punished, and he decides to take the illegal deduction while the device remains idle (Pereboom 2009: 117).

Non-moral Reasons? Why We Won't be Considering Tax Evasion 3 (Much)

Tax Evasion 3 is not easy to grasp and I make no claims to fully understand it. As we shall shortly see, the main difference between Tax Evasion 2 and Tax Evasion 3 is that in the latter case it appears that Pereboom's agent, Joe, may be moved in the direction of paying taxes only by *non-moral* considerations: considerations of self-interest. I am perplexed by Pereboom's change to his example in this regard. We are now considering an agent, Joe₃, who does something by hypothesis immoral (non-payment of tax) and by stipulation will only refrain from doing this for non-moral reasons. This change leaves me struggling to situate the example within the core concerns of the Frankfurt literature⁶.

A Frankfurt argument must minimally establish the conceptual possibility of an agent who is morally responsible (MR) for some act (whether for good—as with character-based counterexamples—or more commonly, ill) yet has no alternative possibilities (AP) open to him. That is, a Frankfurt argument must establish the falsity of the PAP conditional: If MR then AP. Or, put as a truth table, it must establish: i) MR TRUE and ii) AP FALSE. My exegetical difficulties with Tax Evasion 3 may then be drawn out with a dilemma argument. **Dilemma:** either Joe is, as described, not capable of responding to moral reasons (“if he were to refrain from deciding to evade taxes, it would be for self-interested reasons”) or he is capable of responding to moral reasons (“he is aware and sensitive to the moral reasons not to evade taxes in the way he is contemplating”). Both horns of this dilemma appear to have some potential for support by the text.

First Horn of the Dilemma “If he were to refrain from deciding to evade taxes, it would be for self-interested reasons. In fact, it is causally necessary for his failing to decide to evade taxes in this situation that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so...”. Joe, under this disambiguation of the case, is not capable of responding to *moral* reasons, so Joe is not morally responsible; but Joe must be i) morally responsible and ii) lack alternative possibilities. Since i) does not obtain, ii)'s truth (“a mental state he can produce voluntarily”) is irrelevant: the PAP conditional is already true by virtue of its antecedent being false.

⁶ Two core objections shortly to be made against Tax Exemption 2 (“The Basic Response to Tax Evasion 2” and “Question is Begged in Describing the Device as Compelling Joe's / Jones' Choice”) will clearly generalise to other aspects of Tax Exemption 3; so I am not just saying “I don't understand a key aspect of this example” and leaving it at that. Nevertheless, this section's dilemma argument is as much an attempt to bring critical pressure to bear towards clarifying this example as it is an attempt at knock-down refutation of it. Of note is that Pereboom (2012) returns to Tax Exemption 2 (‘moral reasons’) despite, in 2009, having stated that Tax Exemption 3 was his preferred variant of the case.

An agent who is stipulated as being incapable of being moved by any but *non-moral considerations* is in no obvious sense morally responsible⁷—saving certain further, tracing, etc. details and qualifications that have no obvious ability to affect the core issues in this case. Pereboom’s ‘non-moral’ change, from Tax Evasion 2 to Tax Evasion 3, is one that I am then at a loss to understand; unless, which I strongly suspect, in Tax Evasion 3 a prior-stage genuinely *moral* decision-point is tacitly assumed – at which (moral) decision-point Jones is held to have the ability to choose “in a libertarian sense” as to whether to bring before himself the *non-moral* reasons which he knows can move him⁸. Without some such projective exegetical attribution, the role in this case of the passage “it is causally necessary for his [Joe’s] failing to decide to evade taxes [=pay] in this situation that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so, a mental state he can produce voluntarily” is not easy to understand. The interpretation we are entertaining is that, *pace* the ‘non-moral’ claims, we are meant to see Joe as having some highly dissociated, prior, purely *moral* decision-point as to whether to functionally, prosthetically, bring before his imagination the *non-moral* considerations that he realises can alone move him (though *cf.* Mele (1997, 2001) and Lockie (2003)—the static and dynamic paradoxes of self-deception). Might this be what is meant by the otherwise odd claim (in light of the ‘non-moral’ stipulation) that Joe is nevertheless “aware and sensitive to the moral reasons not to evade taxes in the way he is contemplating”?

Were anything like such an exegetical attribution to be correct, the following points would need making: Firstly, not enough of this is in the text, it is a speculative, projective attribution on my part, trying to make sense of an otherwise perplexing example. Secondly, any such interpretation would make this into a highly convoluted case; one vulnerable to criticism for invoking dubiously coherent and oddly stipulative psychological details, of which more below. Nevertheless, were Pereboom to embrace the second horn of the dilemma he might well need to invoke some such prior-stage moral decision-point. We consider this horn of the dilemma now.

Second Horn of the Dilemma Joe is capable of responding to moral reasons. He “has a strong but not always overriding desire to advance his self-interest regardless of its cost to others”. *Not always* overriding, note. Joe, under this disambiguation of the case, “is aware and sensitive to the moral reasons not to evade taxes in the way he is contemplating”. However, “in this case these reasons are overridden by self-interest”. Since, by hypothesis, he *could* have responded to his moral reasons, this is then just what he *ought to have done*. He should have acted in the direction his moral sensibility indicates and *not* have overridden his moral reasons. The case as indicated allows “that he imagine, with a certain degree of vividness, being at least fairly severely punished for doing so, a mental state he can produce voluntarily”. Since he *can* do this, this is what he *ought* to do: “voluntarily” produce that imagining. This is his alternative possibility.

⁷ This is one of the responses Moya (2011) advances against Tax Evasion 3. In advancing this response Moya is building on the argument he first used in Moya (2007) against Widerker—as remarked upon briefly in the section entitled “Decide vs Deliberate” above: that Widerker’s agent Jones could not be seen as morally responsible if Jones was not capable of responding to reasons (reasons to visit his uncle rather than stay with Mary).

⁸ In another case (‘Tax Cut’) Pereboom (2012) similarly envisages such a non-moral, yet (one assumes) prior-stage *moral* decision point: of his agent, Jones, being able to prosthetically bring before his imagination the *non-moral* considerations that he realises can alone move him (his boss finding out his immoral course of action).

Suppose it be objected that Joe is capable of responding to moral reasons in some sense that is sufficient for moral responsibility (“aware and sensitive” etc.) But that here he is, as a matter of fact, just not going to be moved by these reasons: “*in this case* these reasons are overridden by self-interest”. Question: what is the determinant of Joe’s selfish interests not being overridden by his countervailing moral reasons “in this case”? Is this determinant *Joe*? (Joe the agent, Joe *himself*). If it is, then he, Joe, is culpable, as he failed to agent-cause or self-determine that his moral reasons prevailed over his selfish interests—that was his alternative possibility.

What other things (besides Joe the agent) might instead be the determinants of Joe’s selfish interests here prevailing over his countervailing moral reasons? Well, it could be natural law and the big bang that determined this. But to assert that Joe is morally responsible despite natural law and the big bang determining his following an immoral and blameworthy course of action, is a refusal to accept PAP, not an argument against it.

So, what other things (besides Joe the agent and natural law) might determine that Joe’s selfish interests are here unmoved by his countervailing moral reasons? The claim that “these [moral] reasons are *overridden* by self-interest” suggests a ‘balancing beam’ (strongest motive) conception of matters. This can be a component of a classic compatibilist picture of things (which in this context, as just noted, would be question-begging). But such a picture is compatible with both an acausal and (especially) an event-causal theory, whereby at a ‘torn’ point of perfect balance or indifference, chance determines that selfishness wins out. If, in the teeth of the *Mind* and replay arguments you are prepared to maintain that in such a case, Joe nevertheless remains morally responsible then the event-causal indeterminist account of PAP is being defended (and *mutatis mutandis* for other, e.g. acausal libertarian theories). Still no (non-question-begging) defence of moral responsibility yet could-not-have-done-otherwise.

I make no strong claims that this dilemma argument is devastating against Tax Evasion 3 since I am not certain that there is one single way to understand this case. It is meant at least to flag some of the exegetical / interpretative difficulties this case presents. Henceforth, for these exegetical reasons, we will consider Tax Evasion 2, since, *pace* Pereboom (2009) I take this to be the strongest and clearest (though not the most recent) of his cases. Two of the main points to be made against Tax Evasion 2 (in sections entitled “The Basic Response to Tax Evasion 2” and “Question is Begged in Describing the Device as Compelling Joe’s / Jones’ Choice” below) will, I think, clearly generalise to core aspects of Tax Evasion 3; and probably some of the later, more minor points also. These points will be flagged where they occur; but issues of interpretative clarity and the non-moral nature of the example mean that Tax Evasion 3 will not be the main case we work through from this point forward.

Pereboom’s 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, 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: 113).

I have illustrated the salient details of this difficult case with Fig. 1 below.

The Basic Response to Tax Evasion 2⁹

The basic response to this example is, as for so many other such examples, a flicker of freedom response—here concerning Joe's options at decision point 1. As stipulated by this example, Joe is incapable of simply choosing to decide to pay taxes directly—this, as intentional-level description, is not an alternative possibility for him. But at decision point 1 he *is* capable of considering his moral reasons—and does not do so. So Joe is culpable for not considering his moral reasons. In fact, as Moya (2011) notes, this response very probably generalises to Tax Evasion 3; it appears to be a version of the second horn of the dilemma levelled against that case, above.

Pereboom (e.g. 2009: 116, 2012) will demur here on grounds that Joe's *considering his moral reasons* is not the same thing as Joe *deciding to pay his taxes*. A strong response, to be dealt with at length in sections entitled “**Robustness, Intensionality, Right vs Good: A Tension**” and “**The Intensionality of Intentions**” is that with this attempted counter the onus will be on Pereboom to explain how, granted that his argument requires defensively that Joe's *considering his moral reasons* is not the same thing as Joe *deciding to pay his taxes*; nevertheless, as his argument requires offensively, refraining from doing the former makes Joe morally responsible for refraining from doing the latter. This, I think, constitutes an exceedingly pointed challenge for Pereboom; but I shall bracket it for now. An issue to be dealt with here is that we work with the details that are stipulated by the example's author. If considering his moral reasons is stipulated to be all that Joe *can* do, then doing this is all that he may be (positively) *obliged*¹⁰ to do, and not doing this is then all he may be *blamed* for doing—

⁹ This response is advocated by Widerker (2011)—though as a second-string response to Tax Evasion 2 (his primary response I do not completely understand). It is more plainly advocated by Moya (2011).

¹⁰ Actually these matters are complicated—with this a fairly unguarded statement requiring a level of qualification which it can only partly get here. A person may be objectively ‘obliged’ to do that which he cannot do, but this objective and unattainable sense of obligation pertains to the good and not the right, and thereby not to responsibility itself and as such (hence the scare-quotes—these issues will be discussed further below). It is also true that there is a venerable and morally serious tradition in ethics, particularly in certain Christian traditions, of taking obligations as an avowedly unattainable (but action-shaping) ideal. And it is true that one may have an obligation that is currently unattainable but within one's ‘zone of proximal development’. These are large issues and this is the nearest I can get to qualifying them here.

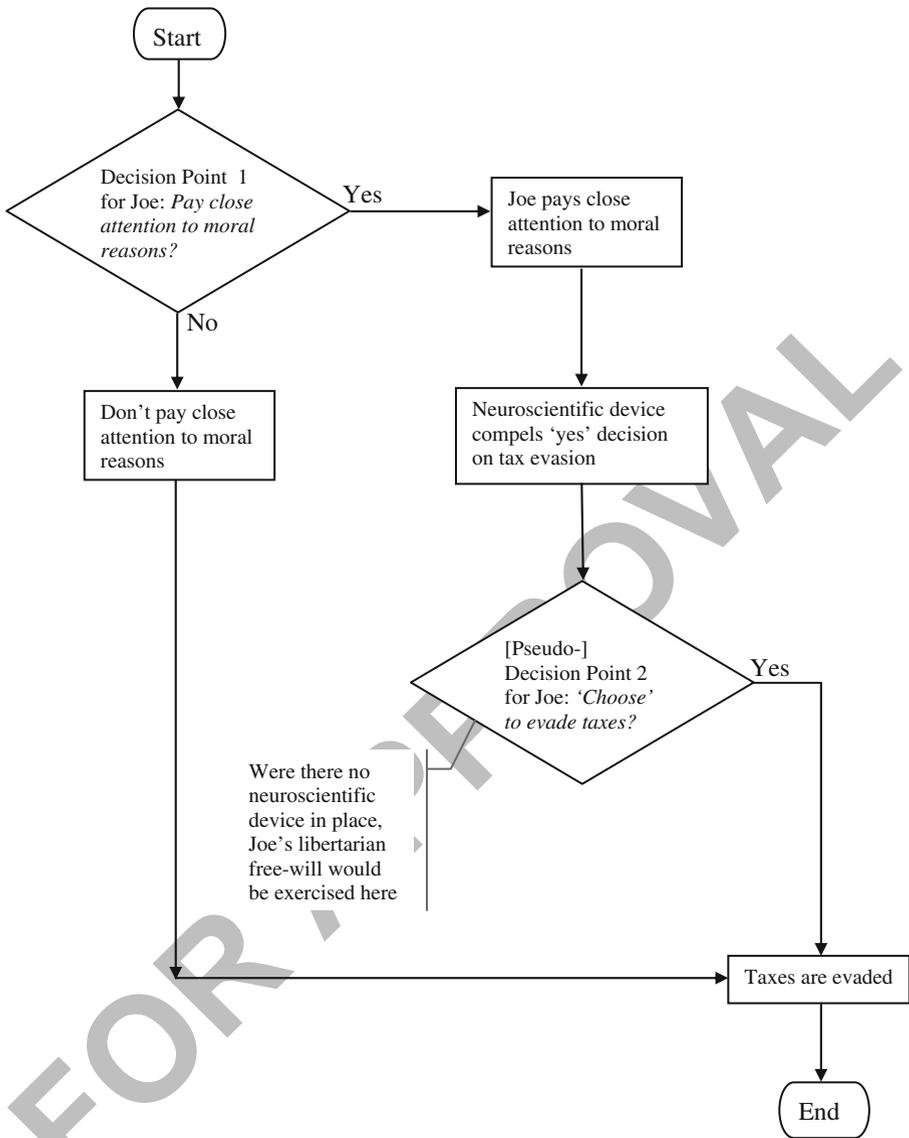


Fig. 1 Pereboom's Tax Evasion 2

all that he may (negatively) be held responsible for: a response in kilter with Otsuka's 'Principle of Avoidable Blame' (PAB), Moya's 'doing everything one can' (DEC), Widerker's 'Principle of Alternative Expectations' (PAE) and Copp's Argument From Fairness.

This statement, from "if considering his moral reasons ..." is undoubtedly *resolute*—and may appear question-begging in invoking PAP in defence of

PAP¹¹. However, several points may be made in response to this criticism even before we get to the section entitled “Robustness, Intensionality, Right vs Good: A Tension” below. One point is that this statement appears to invoke ‘ought’ implies ‘can’ (OIC) not PAP—and OIC is a considerably weaker principle than PAP: “With respect to any action, Kant’s doctrine [OIC] has to do with the agent’s ability to perform *that* action. PAP, on the other hand, concerns his ability to do something *else*” Frankfurt (1983/1988: 95–6). Frankfurt, notably, defended OIC yet rejected PAP¹². I am here rejecting Frankfurt-style arguments against leeway incompatibilism (in part) on the basis of OIC, which Frankfurt, among others, accepted. Here this is, slightly regimented:

- i. Consideration of his moral reasons is all that Joe can do (Pereboom’s stipulation of his thought experiment).
- ii. *Deciding to pay taxes* Joe cannot do. (Corollary of (i)).
- iii. Ought implies can.
- iv. An inability to do φ entails the inapplicability of any obligation to do φ . (iii, contraposition).
- v. Joe’s *deciding not to pay his taxes* is not something he can be blamed for¹³ and Joe *deciding to pay his taxes* is not something here, in this situation (though not in other, counterfactual situations) he may be *required to do*. (Joe can and should be blamed for a lack of consideration of his moral reasons however).

The underlying principles here look in the neighbourhood of Otsuka’s Principle of Avoidable Blame or Copp’s Argument From Fairness or cognate arguments/principles including those adverted to in the section above entitled “Intentional PAP”. So, in advancing the Argument From Fairness, Copp (2003) gives the following example, which I have expressed in numbered premises:

- i. If Jane has the responsibility [=duty, =obligation] to cook a soufflé for everyone in the post office in five minutes [=OUGHT] then she must be able to cook a soufflé for everyone in the post office in five minutes [=CAN].

¹¹ I owe this point to an anonymous referee; and see also Pereboom (2009, 2012)—on ‘dialectically unsatisfying’ responses to his position—of which more below. Fully dealing with this point requires a consideration of the material forthcoming in the section entitled “Robustness, Intensionality, Right vs Good: A Tension”.

¹² A point I owe to an anonymous referee is that Frankfurt may have demurred from this recently (*cf.* Frankfurt 2003: 344). However, whatever Frankfurt’s current view, endorsing OIC whilst rejecting PAP was a view he held, and is a view that Kant, Wolf, and Nelkin have held and do hold. “[B]eing psychologically determined to perform good actions is compatible with deserving praise for them, but .. being psychologically determined to perform bad actions is not compatible with deserving blame” (Wolf 1990: 79), Kant 1933: A547/B575 in Nelkin 2011: 99ff & *passim*). This divergence between OIC and PAP most centrally has to do with issues concerning praise, not blame. Frankfurt’s point, as developed by Wolf (1990) and Nelkin (2011) is that PAP requires us to maintain that doing as we *ought* (being praiseworthy, behaving responsibly—even commendably) requires the ability to do *otherwise*: what Haji (1999) calls ‘dual control’ and McKenna sees as requiring Fischer’s *regulative control*. That we require such control may be called into question (early Frankfurt, Wolf and Nelkin do call it into question). This tradition sees the ‘ought’ implies ‘can’ maxim as saying merely that doing as we ought requires the ability to do as we ought, not to do otherwise than we ought: what Haji (1999) calls ‘singular control’ and McKenna sees as requiring only Fischer’s *guidance control*.

¹³ The point made in the sections entitled “Robustness, Intensionality, Right vs Good: A Tension” & “Intentional Objects and Action in Frankfurt Cases” below is that indeed, this is not in fact something he does—he doesn’t get as far as this decision intensionally and intentionally so-described

- ii. Jane is not able to cook a soufflé for everyone in the post office in five minutes [NOT-CAN].
- iii. So Jane does not have the responsibility [=duty, =obligation] to cook a soufflé for everyone in the post office in five minutes [NOT-UGHT].

A second point against the claim that this argument begs the question is that it may be made by appeal only to a widely endorsed epistemic constraint on responsibility: that, as a principle of justice, we can only be morally responsible for the foreseeable (by us) consequences of our acts (the right) and not to the non-culpably unforeseeable (by us) consequences of these acts¹⁴ (the good)—of which a great deal more below. Though, for certain incompatibilists, such epistemic principles may be motivated by a prior commitment to leeway principles (or perhaps, come all of a piece with such principles) such epistemic constraints on responsibility *are typically defended by compatibilists also*—though on what grounds I cannot say. Pretty much all parties to the free will debates concede important epistemic constraints on responsibility, usually taking these to be a priori.

A third, somewhat related, point, is that the structure of Pereboom’s (and Widerker’s) argument requires that normal moral decision-making may be broken down into a series of discrete, temporally linear stages (whether this is empirically correct is not of concern to us, it is an assumption of these thought experiments). One stage is consideration of moral reasons; the next stage is a full-fledged consideration of whether to evade taxes. Independent of theory, it seems intuitive that Joe can’t be morally culpable for non-payment of taxes until he gets to that stage, and Pereboom has stipulated he does not. If, by stipulation of the thought experiment, he only gets to temporal stage one—consideration of moral reasons—then his actions up to that stage are what he may be held morally responsible for: of which more in “Robustness, Intensionality, Right vs Good: A Tension” below¹⁵.

Finally, and centrally, I wish to reprise with emphasis points already registered in the section entitled “The Metaphilosophical Framework and the Logic of the Dialectic” about the metaphilosophical framework of these debates. In these debates the defender of PAP is engaging in a dialectic to defend against Frankfurt counter-examples directed against a principle he nevertheless holds to be both a priori and fundamental to moral theory. There is a limit to how question-begging a reliance on PAP (or some cognate intentional leeway principle) may be taken to be given that such principles have always been conceived thus by their defenders. Further remarks relevant to this ‘question-begging’ objection are made below—especially in sections entitled “Robustness, Intensionality, Right vs Good: A Tension”, “The Intensionality of Intentions” and in F.N. 22.

¹⁴ We *can* of course, be held morally responsible for culpably unforeseeable consequences of our acts (self-deception, etc.) and were Widerker and Pereboom’s cases to incorporate such features it is likely that their and my intuitions about moral culpability would coincide—but not, I think, to the detriment of PAP or any other leeway principle. *Propos* remarks above it is unclear to me whether Pereboom’s difficult cases feint in the direction of such culpability through wilful ignorance; but were they to do so, this ignorance would be something his agent, Joe, would be responsible for.

¹⁵ I am not committed to any account of what, epistemically, occurs at any given one of these temporal stages—this is for Pereboom / Widerker to stipulate, though if they leave it under-specified then I will try various disambiguations. If Frankfurt theorists invoke temporal stages, then to stipulate a type of epistemic access occurs at one such stage *which is less than Joe deciding not to pay his taxes*, then the agent—here Joe—is *morally responsible* to whichever level of *epistemic engagement* he is stipulated (by such accounts’ authors) to have reached at that stage. More of which below.

Necessary, Sufficient & Counterfactually Uncertain

One major development which came in with this tranche of Frankfurt cases, is that even were Joe to have considered his moral reasons, this would only have been necessary but not sufficient for deciding to pay his taxes—he still might have refrained from the moral course of action at decision point 2 were he (counterfactually) able to do this (were the neurophysiological device not in place to compel this anyway). I do not accept, however, that any convincing line of argument develops out of this. As described, we know that Joe is no plaster saint—he is aretaically so constituted as to be capable of evading taxes. And he does actually commit a moral wrong in failing to consider his moral reasons (probably a considerably worse wrong than the knowing non-payment of his taxes). We convict him of these two things, but the rest is a counterfactual space of epistemic and metaphysical ‘maybes’.

Dummett (1978: 148) makes a cognate point in the context of a different philosophical debate: Jones has led a very sheltered life. He has never encountered danger. Must it be the case that either ‘Jones is brave or Jones is not brave’—where ‘Jones is brave’ means ‘if he had been in a situation of danger he would have acted bravely’ and ‘Jones is not brave’ means ‘if he had been in a situation of danger he would not have acted bravely’? Dummett thinks his realist opponent, to be realist, must affirm this (n.b. I doubt that).

Now, must it be the case that, to judge our putative tax evader to be morally wrong, we have to hold that *had he paid close attention to his moral reasons and the neurological device not been in place, he would have paid his taxes?* I doubt we are committed to that¹⁶. Joe hasn’t done all he can at decision point one and we’ll never know what he would have done later if he had (indeed as non-determinists we can doubt whether there is always a determinate fact of the matter here). But he definitely *hasn’t* done something he should—which is to achieve the requisite level of attentiveness to his moral reasons. For this he is blameworthy, *but not for evading taxes*. I can be known to have done wrong to this, specified, degree, without us knowing whether, had I done otherwise, I would then have done right—or, indeed, done worse still. Is this so different, really, to many situations that obtain in an ordinary life, without appeal to *recherché* cases of this ilk?¹⁷ I am

¹⁶ I should clarify that Pereboom clearly isn’t committed to this either—he holds that had the neurological device not been in place and Joe paid close attention to his moral reasons, he might or might not have then evaded taxes (I owe this point to an anonymous referee). But Pereboom does argue against his leeway opponent that becoming more attentive to moral reasons isn’t a *robust alternative* to deciding to evade taxes precisely on the grounds that one might become more attentive and still not pay one’s taxes. “What is the motivation for thinking that becoming more attentive to the moral reasons now becomes a robust alternative to deciding to evade taxes? Agreed: it is the next best action available to him. But it is not per se exempting for deciding to evade taxes, since, without the device in place, even if he did become more attentive he could decide to evade taxes” (Pereboom 2012: 304). I am arguing here that it is not a requirement on me to oppose Pereboom’s claim that “without the device in place, even if he [Joe] did become more attentive he could decide to evade taxes”—I am claiming that this becomes a counterfactual space of epistemic and metaphysical ‘maybes’ over which the indeterminist of all people does not have to take a stance. I am claiming Joe is culpable for non-attention to his moral reasons, not for non-payment of taxes.

¹⁷ There are very well-worked debates concerning moral luck around this point. Nagel’s German company man who is transferred to Argentina in 1929 is lucky enough to avoid being in a position where he might have become a Nazi collaborator by the 1940’s. In these cases everyday circumstances, not an evil neuroscientist, ensure that we condemn him for, say, ‘tracing case’ aretaic tendencies to conformist ultra-nationalism—but not for what might (possibly, not certainly) have become complicity in murderous acts beyond this. Were one to hypostatise a Dummettian conception of realism to condemn him for counterfactual counterpart time-lines (say, to the point where circumstantial moral luck shades into constitutive moral luck) then “who should ‘scape whipping?’”

blameworthy *now*, to this extent, even though, had I done otherwise and been blameless now, I might yet have been blameworthy later—or not. To be blameworthy now though, there must be other possibilities, of not being blameworthy now: for Joe, the possibility of *paying close attention to moral reasons*; for Jones, *trying to visit uncle*, or at least, *starting to take steps that leave open the possibility of visiting uncle: deliberating in that direction*.

Question is Begged in Describing the Device as Compelling Joe's / Jones' Choice

We are told across both Tax-Evasion-2 and Tax-Evasion-3 that the neuroscientist has implanted a device which, were it to sense the requisite level of attentiveness, would electronically stimulate the right neural centres “so as to inevitably result in his [Joe] making this choice” (Pereboom 2009: 113 (Tax-Evasion-2), 117 (Tax-Evasion-3)). However, conceptually, not empirically, this is impossible—as it's then no longer *Joe's choice*, it's the neuroscientist's—a point first made by Lucas.

‘I did not shoot him. I did not pull the trigger. My muscles tightened in response to electrical stimulation.’ Here we are partly rejecting the *did*—it was not a deed I did but something that happened to me. We are also partly rejecting the *you*—it was not I who had a hand in it (Lucas 1970: 6).

So, the last box in Fig. 1 above reads ‘taxes are evaded’, but with this a statement of consequences, not action¹⁸. The box doesn't read ‘Joe evades taxes’; it doesn't even read ‘Joe culpably ignores his moral reasons, leading to taxes being evaded’, as Joe might not have ignored his moral reasons and still this box would have been reached by another non-Joe route (via the neurophysiological device). This means that the decision point 2 Joe faces—as to whether to evade taxes—is a pseudo-decision point, not a *decision* point at all, as by hypothesis the ‘decision’ in question was a compelled mechanistic output of the neuroscientist's device.

Widerker, in his Brain-Malfunction-W case, also faces this as an objection, an objection he seeks to finesse by maintaining that his agent, Jones, unlike Pereboom's Joe, is not compelled to ‘decide’ (for Jones: to meet with Mary); he just has the alternative option—of *not meeting with Mary* (of visiting his uncle)—removed. So the equivalent of the ‘No’ option at Pereboom's decision point 2 is removed, but unlike for Pereboom's case, no direct neuronal intervention within the decision-process compels Widerker's agent, Jones, to make the choice ‘Yes’. Jones has the option of continuing indefinitely in futile deliberation. Widerker even tries to make this difference between his case and Pereboom's into a critical point against his PAP opponent. With his would-be moral agent unable to do other than repeatedly fail to reach his otherwise preferred, moral, decision (to visit Uncle), Widerker notes that in continuing to attempt to reach the moral (Uncle-visiting) outcome, while yet being neurologically thwarted

¹⁸ See remarks on the right versus the good and the intensionality of intentions. We knew already, long before the Frankfurt debates, that consequentialism has tended, for good reason, to be the moral theory of determinism, as deontology has tended to be the moral theory of libertarianism. Forty-five years of Frankfurt arguments must establish more than that were we unable to do otherwise, we might yet be ‘responsible’ for a given level of consequentialist outcome (the good).

Jones cannot fail to realize that (in the circumstances) continuing to deliberate would mean to decide not to keep the promise. Hence, Jones cannot but view the option of continuing to deliberate and that of deciding not to keep the promise [to visit his uncle] as one and the same (Widerker 2011: 270—& *c.f.* Pereboom 2012: 311).

I think though, that Widerker's adjustment does nothing to deflect the critical objection made after Lucas—that what happens thereafter is no longer *Jones'* decision (or at least, no longer his undamaged and free decision, no longer a decision *he* may be held responsible for). At an intentional-level description (at the level of the right not the good) Jones is not aiming to cogitate in a circle (see points under “[The Intensionality of Intentions](#)” section below). He's trying to do the right thing—but a neurological stymie is preventing him. Pointing out that he may be stymied for so long that his option of effecting the good is lost to him (and he sees it is lost to him, and this affects his eventual action) does nothing to counter the Lucas objection. Jones may be mutilated, neurologically, so badly that for a given decision no moral appraisal is possible of him; but by hypothesis this is supposed not to be such a case (any Frankfurt example needs its agent to i) *possess* moral responsibility and ii) *lack* AP). Widerker needs Jones to be sufficiently free from Zombie, Skinner-box, etc. mutilation to permit us to evaluate him (harshly) for choosing of his own free will to cavort with Mary rather than do all he is able to try to visit his uncle. If he is able to seek to do what he *ought*, though this, alas, leads him to sit and spin in neurologically thwarted cogitation as he relentlessly butts his head against stymie -*N*, then he has done all he *can*. If, after spinning for a while, he (bewildered by his sudden-onset apraxia) reels away and into the arms of Mary, then I for one can't find it in my heart to condemn him.

Problems with Accepting the Terms of These Examples

There is a lot going on at the same time in these examples. Among the problems with accepting these examples on their own terms are the following.

What Joe is Responsible for

Pereboom wants the judgement of moral irresponsibility against Joe to be that he has evaded taxes. Widerker (2011) and Moya (2011) offer us the clearly available judgement against Joe that he has irresponsibly failed to consider his moral reasons sufficiently. Pereboom feels this is an insufficiently robust judgement. Insufficiently robust for what? We attribute to Joe the responsibility appropriate to the details of his case—as Pereboom was free to stipulate these. If our judgement is insufficiently robust to credit Joe with, for example, his being epistemically apprised of the fact that “I, Joe, am evading taxes and that is morally wrong” then this is because Pereboom expressly doesn't give us this. He gives us an awkward oscillation towards and away from it: *I, Joe, am evading taxes and that is illegal. I, Joe, am capable of not evading taxes but only by considering deeply my moral [Joe₃= non-moral] reasons. I, Joe, am [culpably] not considering my moral [Joe₃= non-moral] reasons.*

Joe is knowingly committing an *illegal* act—not that that’s relevant to our moral case. And morally, not legally, he’s culpably avoiding consideration of his moral reasons. And it’s metaphysically, not just epistemologically, uncertain whether he would have paid his taxes were he to have discharged his moral responsibility to consider his moral reasons. What then do we say of him? Well, we say he’s culpably immoral for not considering his moral reasons—and that’s bad to the extent it’s bad (probably a rather greater extent than knowingly and with malice aforethought not paying his taxes). We also note that Joe is capable of some fairly sophisticated levels of active self-deception around moral issues, which is aretaically a poor show; and also that Joe has an odd, highly dissociated psychology. But, as described, we should hold he is morally responsible for wilful inattention to his moral reasons, and legally responsible for non-payment of taxes, and aretaically he’s highly suspect around issues of self-and-other deception; and that’s as far as we go. If Pereboom wants a more robust judgement against Joe than Moya’s ‘culpable failure to consider his moral reasons’ he must state the psychological details required to warrant our drawing that more robust judgement. The case as described warrants the judgement as described. It could be disambiguated otherwise and warrant another judgement.

To make vivid some of the points made here, let us alter the description of this case in one respect only: make the act much more morally grave than evasion of taxes but keep all else the same.

Moral Triviality and Moral Gravity

There are sins of omission as well as commission. One point to note is the moral triviality of the wrong our revenue-avoidant businessman is putatively guilty of. What work is that doing in the example?

Saving orphans from slavery and murder evasion. A businessman has, by chance, discovered that a murder has been commissioned, in his interests but unsolicited by him—commissioned by a business associate in order to curry favour with him. Let’s say, it is the murder of a priest who runs an orphanage (the location of said priest’s orphanage is an obstruction to that businessman’s completion of a bit of land speculation). The businessman is minded to avoid cancelling this murder; with the only possibility of that businessman’s non-evasion (of cancelling his acquaintance’s murder-contract) is if that businessman pays close attention to his moral reasons—but he doesn’t, the priest is gunned down, and the orphans sold into particularly nasty forms of slavery.

Here, we surely don’t think there is any excuse, either on grounds that the businessman didn’t pay close attention to his moral reasons; nor that he needs to take action to prevent murder only in this perversely dog-legged way—*sub specie* “consideration of moral reasons” in the abstract, rather than concrete moral horror at this particular act. If that businessman were to start to consider his moral reasons, we suppose a neuroscientist would thwart him. If the neuroscientist didn’t thwart him he might still decide to let the murder go ahead. What

should we make of this? Surely that he's a vile human being for being capable of correcting this moral infamy only in this dog-legged way: *sub specie* "consideration of moral reasons" in the abstract, rather than concrete moral horror at this particular act; and moreover that he could have simply repudiated the murder immediately—that was his choice.

A Very Convoluted Case: Dissociated and Stipulative Psychological Details

The foregoing two sub-sections raise a matter that is insufficiently discussed in the copious Frankfurt literature. This is the oddly dissociated psychological details of so many of these examples. Pereboom's case, as described, gives with one hand and takes away with another—in a way that is crab-like, stipulative and obscure. It clouds our ability to reach any simple judgement. It is also exceedingly difficult to understand. We are told: "The only way that in this situation he [Joe] could fail to choose to evade [=pay] taxes is for moral reasons, of which he is aware". Of what is Joe aware? Directly: that is he immoral? The whole point of needing the example to be this oblique militates against that straightforward a reading. Is it anything like the following? Joe has a kind of dissociative, prosthetic, third-person knowledge-of (and yet immediate control-of) his own psychology, such that:

- i. Joe has transparent, explicit, occurrent, knowledge that he is capable only of paying taxes for moral [Joe₃=non-moral] reasons; yet
- ii. Joe has absolutely no epistemic registration of the fact that these reasons indicate (morally, as opposed to legally) that he *should* pay taxes (this on pain of him being simply a culpable, knowing, tax evader); and
- iii. Joe has some motivated reason to not "attain a certain level of attentiveness to these reasons" which might permit disclosure of ii); and
- iv. Joe has the (stipulated) perfect, immediate, executive ability to act on that reason ("exercising his libertarian free will"—whatever that means: pressing a free-will button?)

If this is the picture we are operating with, is it psychologically plausible? It is at least odd, contorted, and psychologically under-described. But it may be rather more than that. For reasons I indicate at length in Lockie (2003) I think this type of reasoning is conceptually confused. For Joe to commit a motivated *act* of culpable self-non-disclosure it is required that at some level he have registered the fact that he is seeking to defend against (*pace* Mele (1997, 2001) he needs to have seen 'both sides of the limit': the static and dynamic paradoxes of self-deception). And that just convicts him of morally culpable self-deception in the service of already (at some level) registered-as-immoral ends. He could have reasoned more honestly—without said self-deception. Paying more attention to his moral reasons is precisely what the discharge of his intellectual (and moral) responsibilities would here consist in.

What Kind of Libertarianism is Conceded in These Examples?

A feature shared by both Widerker's and Pereboom's examples is that they stipulatively concede to their PAP opponent a species of libertarianism. However, *what* species of

libertarianism is problematic. The decision of Widerker's Jones is "to be understood in a libertarian sense, implying that such a decision is neither nomically determined nor caused" (Widerker 2011: 269). That looks to be a simple acausal indeterminism, a position which, among current libertarians is, with good (*Mind* argument) reason, rather a minority taste. The agent-causal libertarian, for example, is highly unlikely to be satisfied with this. Nor are such libertarians likely to be any more satisfied with Pereboom's claim of his agent, Joe, that "he could, exercising his libertarian free will, either choose to evade taxes or refrain from so choosing" (Pereboom 2009: 113)¹⁹. This sounds like a kind of 'frictionless' freedom of indifference or caprice, whereby, against Pereboom's earlier stipulation, Joe indeed "could choose to evade taxes for no reason or simply on a whim" (Pereboom 2009: 113). What is the PAP libertarianism, which Pereboom and Widerker separately oppose, envisaged as amounting to?

Widerker sees himself as a source indeterminist; though with a view which is explicitly not agent-causal (2009: 101). I myself take it that we precisely need a metaphysically committed notion of the agent/self as a source in order to be *source* libertarians, but that view is undoubtedly controversial. Widerker (2009:101); very explicitly assimilates his (he claims, source indeterminist) view to *non-causal* libertarianism; and this I take to be specifically problematic. From the early modern period onwards, the stock *Mind* objection surfaces: how can the agent be the *source* of her act, sufficient for moral responsibility to be predicated of her, if the act is random, by chance, acausal, (positively) undetermined? Whether you think the *Mind* dilemma a good objection or not, it is precisely the objection that the agent can't be the source of her actions if randomness is. Positive indeterminism of this acausal type embraces one horn of the *Mind* dilemma, it doesn't adopt the position that this is a false dilemma—say, that the agent or self is instead the source.

The Frankfurt argument is meant to be a general, conceptual-level argument against AP being a necessary condition on moral responsibility. Widerker and Pereboom are each conceding a libertarian freedom to their PAP opponents in their examples (in order not to beg the question against these opponents—which would be one horn of the WKG dilemma). But the species of libertarian freedom they concede matters a great deal, and cannot be stipulated merely as a cipher: 'libertarianism' per se. It is dubious whether these arguments will work for each and every species of libertarianism that

¹⁹ Pereboom, though a hard incompatibilist, is explicit across many publications that he regards agent-causal libertarianism as conceptually possible and, were it actual, as capable of underwriting a responsibility-relevant free will—so unlike, say, Galen Strawson, he's no 'impossibilist'. However, he has hitherto regarded the agent-causal position as untenable given our current best physical theories (though there are intriguing recent signs he may be changing his views positively in the direction of agent-causation). Various agent-causal theorists do see their theories as committing them to PAP however—so they will feel that the Pereboom thought experiments' specific conceded libertarianism does not fix an argument that applies against their position. There is reason to believe Pereboom has in his sights an event-causal libertarianism; but note that Pereboom in his (2001) appears to have been using 'event-causal' in a sense that now we would understand as 'acausal'. Certainly the language Pereboom uses in 2009 and 2011 to describe the freedom of his agent, Joe, has clear aspects of acausal libertarianism and also aspects of an older, less specific 'freedom of indifference' (the classic compatibilists' nowadays more-or-less straw-man libertarianism). Widerker is an acausal teleological-intelligibility libertarian but he acknowledges (Widerker 2009: 101) that he can envisage "other versions of Frankfurt-friendly Libertarianism" (his term for a source-and-not-leeway position). The question for Widerker and Pereboom is not whether they can envisage other species of libertarianism (clearly they can and do) but whether their attempted conceptual-level refutations will work against any and every notion of PAP libertarianism their *opponents* envisage—and advocate.

may be substituted into their stipulative libertarian clauses; indeed, they may tacitly be working with a highly proprietary notion; only one out of the PAP libertarianisms that are available²⁰.

Robustness, Intensionality, Right vs Good: A Tension

‘Robustness’ applies to alternative possibilities. Pereboom anchors his notion of this in epistemic awareness. I have italicised the epistemically relevant phrases in the two quotes that follow (we will not be concerned with the analytic details). In ‘Robustness A’ Pereboom notes:

For agent to have a robust alternative to her action A, that is, an alternative relevant per se to explaining why she is morally responsible for A, *she must understand* that instead she could have voluntarily done something as a result of which the specific sort of moral responsibility she actually has for A would have been precluded (Pereboom 2001: 26, 2012: 300).

Pereboom reformulated this in the teeth of detailed analytic challenges from Nelkin and others, into ‘Robustness B’, whereby the epistemic clause requires of the agent

that for at least one such acting or refraining, she *is cognitively sensitive* to its [the alternative to her action] *being available to her*, with the result that *she believes to some significant degree* that had she voluntarily so acted or refrained she would be, or would likely be, blameless (Pereboom 2012: 301).

Widerker, Pereboom et al. claim that their opponents’ alternative possibility is not sufficient to count as a *robust* alternative to the act in question (uncle non-visiting, tax avoidance) because it lacks more than a certain (call-it ‘level one’) epistemic access status (more than deliberation [q.v.] (Widerker), more than ‘voluntarily achieving the requisite level of attentiveness’ (Pereboom)). This non-robust, level-one alternative does not go all the way to (call-it ‘level two’²¹) epistemic access: Jones *deciding* [q.v.] (Widerker). Joe “has *no inkling, and no evidence*, that the [neuroscientist’s] intervention would then take place, as a result of which he would not be blameworthy” (Pereboom 2012: 303). In other words, the agent does “not .. *understand* that instead she could have voluntarily done something as a result of which the specific sort of moral responsibility she actually has for A would have been precluded”; the agent is not “*cognitively sensitive* to its [the alternative to her action] *being available*.”

²⁰ To be blunt: there is a suspicion that the core target of this argument is simple positive indeterminism (whether acausal or event-causal, but substantially unanalysed beyond being positively undetermined: the classic ‘freedom of caprice’ or ‘indifference’—a conception of liberty that now borders on the straw-man). Any such indeterminism is already, surely, well refuted by the *Mind* argument. A sophisticated modern leeway incompatibilist (who may *also* be a source incompatibilist) may be an agent-causal theorist or a complex event-causalist or anything else. Any such theory will hold that a specific sub-set of *morally relevant* ‘alternative possibilities’ are (sometimes) *necessary* for freedom and/or moral responsibility; not just any old non-moral alternative possibilities are *necessary and sufficient* for such.

²¹ If this ‘levels’ terminology obscures rather than assists understanding, then please assimilate the points made without it.

But Pereboom & Widerker must and do insist that this non-robust ‘level one’ epistemic access *is* sufficient for responsibility-relevant moral appraisal of the act in question *under the full-fledged intentional/intensional description* (uncle non-visiting, tax avoidance). So, in the last sentence of a crucial paragraph emphasising the above, negative, point—the lack of a *robust* alternative due to his agent, Joe’s, epistemic unawareness—Pereboom (2012: 304) says “Nevertheless, intuitively Joe is blameworthy for actually so deciding [to evade taxes]”. And I just don’t see how that is in any way a defensible position. Of course one may mark differences and nuances between two or many levels of epistemic access/moral appraisal. *But these levels travel the same way for robustness of intensional-intentional representation of alternative possibilities and appraisal of moral responsibility.* And this is for an important reason. There is a categorical difference between the right and the good. The former is an intensional (and intentional)-level kind of moral appraisal. The latter is an extensional, causal, consequential species of moral appraisal. When it comes to moral responsibility—true responsibility, responsibility properly so-called—we are and must be operating with an intensional (and intentional) level of evaluation: with the right and not the good. And that means there can be no space to insert a divergence between the epistemic, subjective-awareness, intensional (and intentional)-level of evaluation and the moral appraisal of the act. *Of course* there is such a space when we switch to a discussion of the good—a well-meaning act unintentionally leading to foul consequences, say; or vice versa. But the good per se is *irrelevant* to moral responsibility, unless we are punning on ‘responsible’ (‘causal versus moral responsibility, etc.) or smearing the debate or otherwise changing the subject (say, by establishing an avowed moral responsibility under an intensional description, then smoothly moving to talk of moral responsibility for the extensional equivalent / causal consequent that is thereby brought to pass). Saving tangential, digressive details (culpable inattention to the likely consequences of their acts, etc.) Joe and Jones are as *morally responsible* (properly so-called) as they are *epistemically aware*. Our intentional-level moral appraisal of them (the right) must travel in the same direction and to just the same extent as their subjective epistemic awareness, intensionally so-described²². Any sophisticated nuances we draw

²² Pereboom considers and rejects the view that Joe is only responsible for culpable inattention to his moral reasons, not tax avoidance, maintaining that “this is a dialectically unsatisfying response to a Frankfurt example, for the reason that it explicitly cites a leeway position in support of its verdict ... Joe is non-derivatively morally responsible only for not deciding to be more attentive to the moral reasons, for only relative to this decision is a robust alternative available to him” Pereboom (2012: 306). In addition to the points made in “The Basic Response to Tax Evasion 2” section above, note: a) everyone in such examples is arguing from their intuitive judgements about moral responsibility: Pereboom is and so are his opponents. I have no intuitive judgement that in these cases Joe is morally responsible, intensionally so-described for *tax avoidance*—but see remarks on the confusing nature of these cases above. b) In drawing judgements as to *responsibility* it matters a great deal that we be precise in these judgements. Such judgements (and not just in the specialised Frankfurt literature) are drawn at the intensional and intentional level, for it is at this level that we appraise people for the *right* rather than the *good*—and it is a basic principle of justice (e.g. Copp’s ‘argument from fairness’, Otsuka’s PAB—or a cognate principle) that we do so. Pereboom (2001) himself bases his hard determinism most generally on such principles (that it would be *unfair* to blame people in a determined world). c) Pereboom’s case is meant to reanimate the anti-PAP literature to victory not stalemate. d) Pereboom himself argues for *non-robustness* (for him, a crucial technical term/device) on the basis of the epistemic opacity (intentional and intensional) of the agent’s actional outcome to himself: Joe cannot *foresee* that his action would absolve him of moral responsibility. For all these reasons I think the claim that this response is dialectically unsatisfying is itself dialectically unsatisfying.

epistemologically must be drawn in the same direction as they are ethically, *provided we are concerned with the right and not the good*—and it is the former that pertains to issues of moral responsibility, properly so-called. We can't be fully responsible for *tax evasion* intensionally so-described if we are epistemically aware (“cognitively sensitive”) only to the fact of our culpably not ‘voluntarily achieving the requisite level of attentiveness’ to our moral reasons. If the latter is held not to represent a *robust alternative* because it doesn't come anywhere near an epistemically transparent representation of how we may achieve the former, then what Joe is *responsible for* cannot be culpable non-payment of taxes, properly so-called. The search for other, even more sophisticated thought experiments, putatively to establish such a divergence (between, on the one hand, a level of intensional-intentional epistemic access to what one is doing that is too superficial to be a *robust alternative*, yet, on the other hand, a full-fledged moral appraisal of the act in question *under said intensional description*) then appears futile. These two things are internally connected because responsibility, properly so-called (i.e. without ‘punning’ or smearing, etc.) is a species of moral appraisal that occurs only at the level of the right and not the good.

Pereboom seeks to rebut the Moya challenge here

The Frankfurt defender can agree that Joe is blameworthy for not becoming more attentive to the moral reasons, and that for this he does have a robust, exempting alternative possibility. But it's intuitive that Joe is also blameworthy for deciding to evade taxes, and for this, at least *prima facie*, he has no robust alternative. What is the motivation for thinking that becoming more attentive to the moral reasons now becomes a robust alternative to deciding to evade taxes? Agreed: it is the next best action available to him. But it is not *per se* exempting for deciding to evade taxes, since, without the device in place, even if he did become more attentive he could decide to evade taxes, and, we might suppose, he even would be likely so to decide. And he is not at all aware that due to the device becoming more attentive would preclude him from deciding to evade taxes, and would thus render him blameless for this decision (Pereboom 2012: 304).

Certain critical points need making in response to this. Firstly, it is not “intuitive that Joe is also blameworthy for deciding to evade taxes”—I have argued above that it is not even true. Secondly, in response to Pereboom's “it is not *per se* exempting for deciding to evade taxes, since, without the device in place, even if he did become more attentive he could decide to evade taxes”: for the clause “even *if he did* ... he *could* ...” see my remarks on counterfactuals and Dummett cases, above. And for issues pertaining to talk of “*exempting*...” this claim may perhaps have force against Moya—or perhaps not—but it does not have force against one who is not seeking to claim Joe may act to make himself *exempt for trying to evade taxes*, arguing instead that Joe, as intensionally described by Pereboom's thought experiment, is not trying to evade taxes anyway. Joe is wallowing in a hard-to-pin down, oscillating and obscure state of intellectual and/or moral dishonesty, which does not possess robust representation as tax avoidance to him (though it veers obliquely towards and away from this) and thereby does not need and cannot be given an alternative course of action that would represent ‘exemption for tax avoidance’—for this latter is not clearly what Joe has done. Joe *could*, were he to follow the alternative course of action that is available to him, be given ‘exemption for

culpable inattention to his moral reasons’—because that is what he *has*, as described, irresponsibly done; and also because that is what he *has* as described, got a robust (*intentional, intensional, epistemically pellucid*) alternative possibility for.

Intentional Objects and Action in Frankfurt Cases

The Intensionality of Intentions

Mid-way through Chisholm’s most famous paper on free will he reacts to an anticipated objection which he develops out of his frontispiece quotation, illustrative of his (and Aristotle’s) agent-causal libertarianism: “The stick moves the stone and is moved by the hand, which is moved by the man” (Aristotle 1999: 256a 6). Chisholm’s anticipated objection to his own position involves noting that the brain must be involved in causing the hand to move, yet he notes (first making his anticipated opponent’s objection in quotes):

“If the *man* does anything, then, as Aristotle’s remark suggests, what he does is to move the *hand*. But he certainly does not *do* anything to his brain—he may not even know that he *has* a brain” ...

[Chisholm responds] The answer to this objection, I think, is this: It is true that the agent does not *do* anything, with his brain, or to his brain, in the sense in which he *does* something with his hand and does something to the staff. But from this it does not follow that the agent was not the immanent cause of something that happened within his brain (Chisholm 1964 in Watson 2003: 31).

Chisholm’s Aristotelian agent has the power to move a stone (which alone is his *intention*—the *object* of his *act*). He does this by moving his hand, which moves his staff. He doesn’t take as his intentional object the movement of synaptic vesicles across certain synaptic clefts towards certain dendrites. He doesn’t even know he possesses these items or is effecting these processes (*cf.* Smart 1959: Objection 1). These neurological events are occurrences that happen as a result of his intention to move the stone; occurrences which (in a harmless sense) *he*, the agent, may be said to have brought to pass through his avowed power to intentionally move the stone, by moving his staff, by moving his hand.

Consider this in light of the way very many pro-Frankfurt philosophers gloss their ‘flicker’ opponents’ positions, the better to make their various ‘insufficiently robust’ objections to these positions (here, say, in regard to a classic Frankfurt case). The flickering agent, *pace* the ‘insufficiently robust’ objection, is not held by the libertarian flicker theorist to have the meagre power *to make a prior sign* (to blush, for example). Nor is she held by the libertarian flicker theorist to have the strange power *to make the neuroscientist intervene*. She is held to have the power *not to kill*. If she were not going to kill, then (highly dubiously in light of the WKG dilemma) the omniscient neuroscientist is supposed able to detect this with perfect accuracy from said prior sign, and himself effect the killing—perhaps by producing in the hapless agent the movement of synaptic vesicles across certain synaptic clefts towards certain dendrites. It is no objection here to say that the agent had *no intention to blush* or *no awareness of the*

intervener, or *no awareness she had neurons* hence (a fortiori) *no intention to produce action by that intervener*, or *movement across those synapses*, etc. Of course the flickering agent had no intention to do these things—who said she did? The flicker theorist maintains (rightly) that the agent is precisely manipulated because Intervener saw she could have been about to develop *no intention to kill*. Indeed, had the agent been permitted to continue undisrupted all the way to a mental precursor to action, she could²³ have formed (stronger) the *intention not to kill*. The fact that she had no *power to prevent Victim's killing* [a statement of consequences] is irrelevant to rebutting the obvious power she did have: *not to kill*. This is just the Lucas point raised in the section entitled “Question is Begged in Describing the Device as Compelling Joe’s / Jones’ Choice” above.

So, when, in response to Widerker, we maintain that *Jones is morally responsible for not butting his head against neurological stymie -N*, we are certainly not maintaining that he knows that that is all he can do, or that that is, or ought to be, his intentional aim. In appraising a putatively responsible agent for whether he has done as he ought, that agent isn’t to be evaluated for the extensional but unforeseen consequences of his actions (the good), he is to be evaluated for the intensional and intentional actional goal (the right). Intentional contexts are intensional contexts (Lockie 2003; Moya 2006: 68ff). As argued above (e.g. in sections entitled “The Basic Response to Tax Evasion 2”, “Robustness, Intensionality, Right vs Good: A Tension” and in F.N. 22) Joe is plausibly only morally responsible (here, in this trivial case, though not *Moral Triviality and Moral Gravity’s* ‘murder evasion’ analogical extension) for not considering his moral reasons. He *might* be described more fully, in a different thought experiment, as having partial epistemic access (through some chain of self-deception say) to his tax evasion, properly so-called—a tax evasion occurring in part through and by a lack of consideration of his moral reasons. Such evasion might be more dishonest and morally more wrong even than straightforward knowing tax evasion, due to the plausible claim that culpable intellectual dishonesty and inattention to moral reasons are worse than knowing tax avoidance. But were any such case to be constructed, it would not alter the fact that there are epistemic limits to moral responsibility; that we have responsibility, properly so-called, within said epistemic limits, that (*salve* issues of culpable intellectual dishonesty, tracing cases and the like) responsibility and awareness travel in the same direction.

Self-Absorption and the Aims of Responsible Conduct

We saw above, Pereboom’s changing attempts to characterise what a ‘robust’ alternative possibility is, with talks of his agent being responsible in case “she could have willed something other than what she actually willed such that she understood that by willing it she thereby would have been precluded from the moral responsibility she actually has for the action” (Pereboom 2009: 110, citing Pereboom 2000, 2001: 26). To

²³ Were we dealing with versions of these arguments developed prior to Frankfurters carefully seeking to obviate the WKG dilemma, both this ‘could’ and that of the last sentence would have been ‘would’s. Note that, in light of the need of the Frankfurters to de-fang the ‘question-begging’ horn of that dilemma, even the later cases’ Frankfurt agents do have to be credited with possessing the strong libertarian power stated in the text (“Question is Begged in Describing the Device as Compelling Joe’s / Jones’ Choice” section after Lucas): the power *not to kill*.

this, a counter immediately suggests itself: at an *intentional* [/intensional] level, Pereboom's agent ought not to be attempting to do this²⁴. There is actually an identifiable *vice* (a vice very much of our times) associated with the agent trying to be "precluded from the moral responsibility she actually has for the action": Think of 'defensive medicine', think of modern 'corporation man' and the efforts of those putatively with responsibility trying to memo their way to a diffusion of that responsibility. The morally responsible agent should be trying to *do good and avoid harm*. There is nothing it would be to aim after "being precluded from .. moral responsibility" direct—nothing that represents a *moral* course of action, in any event. Pereboom subsequently goes through a series of revisions of his characterisation of robustness (the latest of which is 'Robustness 3') without, it seems to me, ever escaping this point: his agent needs to be aiming to *do good* not to *be right*. Even by 'Robustness 3' his agent is still characterised as morally responsible to the extent that she could have tried to be "precluded from the responsibility she actually has" (Pereboom 2009: 112). A moral agent should not be concerned with preclusion from responsibility: she should be concerned to do what she ought to do as best she is able to apprehend this.

Conclusion

I conclude that three of the most recent, state-of-the-art, Frankfurt cases, cases designed to circumvent previous generations of anti-Frankfurt objections (including, notably, those highlighted in the first part of this paper) fail in their aims. To go further than this, and maintain that any such arguments will fail, or must fail, is far more than I can claim to have established in this paper. Nevertheless, certain themes running through this piece constitute, I believe, standing obstacles to future lines of development in the pro-Frankfurt literature. These include a number of intertwined points dealing with the intentionality of intentions; with responsibility properly-so-called applying at the level of the right and not the good; with related issues of deontology vs consequentialism; with there being epistemic constraints on responsibility properly-so-called; with moral responsibility having to travel in the same direction and (waiving qualificatory details) to the same extent as epistemic access; and with the libertarian not being required to nail down 'counterfactual maybes'. Other critical points concerning the complexity, equivocation and ambiguity of these cases have been registered. Many of analytic philosophy's debates are nowadays exceedingly focussed—at times, and in the case of certain debates, perhaps rather too focussed. I ask my reader's forgiveness if, in this regard, I too have "contributed my share" (Frankfurt 1988: 117)²⁵.

²⁴ Despite being in sympathy with Otsuka's (1998) position generally, I would, for these reasons be wary of embracing his choice of terminology ('PAB': principle of avoidable *blame*).

²⁵ I should like to record my gratitude to Carlos Moya who generously made his (then) unpublished work available for me to read and whose philosophical influence (and in many cases, priority) should be clear in this paper. I should also like to record my gratitude to Dana Nelkin for reading and providing generous comments on an earlier version of this paper. Finally, I should like to record my gratitude to an anonymous referee for this journal for acute, incisive comments on earlier drafts.

References

- Aristotle. (1999). *Physics* Book VIII, Clarendon Aristotle Series, tr. Daniel W. Graham, Oxford: Clarendon Press.
- Bealer, G. (1996a). A priori knowledge and the scope of philosophy. *Philosophical Studies*, 81, 121–142.
- Bealer, G. (1996b). On the possibility of philosophical knowledge, *Noûs*, vol. 30, Supplement: *Philosophical Perspectives*, 10, Metaphysics 1–34.
- Chisholm, R. (1964). Human freedom and the self. In G. Watson (Ed.), *Free Will*, Oxford Readings in Philosophy 2nd edn. Oxford: Oxford University Press 2003: 3–15.
- Copp, D. (2003). “Ought” implies “can”, blameworthiness, and alternative possibilities. In D. Widerker & M. McKenna (Eds.), *Freedom, responsibility and agency: Essays on the importance of alternative possibilities* (pp. 265–299). Aldershot: Ashgate Press.
- Dummett, M. (1978/1963). ‘Realism’, in his *Truth and Other Enigmas* Cambridge, MA: Harvard University Press 145–65.
- Frankfurt, H. (1969). ‘Alternate possibilities and moral responsibility’; in his (1988).
- Frankfurt, H. (1983). ‘What we are morally responsible for’; in his (1988).
- Frankfurt, H. (1988). *The importance of what we care about*. Cambridge: CUP.
- Frankfurt, H. (2003). Some thoughts concerning PAP. In D. Widerker & M. McKenna (Eds.), *Freedom, responsibility and agency: Essays on the importance of alternative possibilities* (pp. 339–345). Aldershot: Ashgate Press.
- Ginet, C. (1996). In defense of the principle of alternative possibilities: Why I don’t find Frankfurt’s argument convincing. *Philosophical Perspectives*, 10, 403–417.
- Haji, I. (1999). Moral anchors and control. *Canadian Journal of Philosophy*, 29(2), 175–203.
- Kane, R. (1996). *The significance of free will*. Oxford: Oxford University Press.
- Kant, I. (1933). *Critique of pure reason*, tr. N. Kemp-Smith, Basingstoke Hants & London: Macmillan.
- Lockie, R. (2003). Depth psychology and self-deception. *Philosophical Psychology*, 16(1), 127–148.
- Lucas, J. R. (1970). *The freedom of the will*. Oxford: Oxford University Press.
- Mele, A. (1997). Real self-deception. *Behavioral and Brain Sciences*, 20, 91–136.
- Mele, A. (2001). *Self-deception unmasked*. Princeton: Princeton University Press.
- Moya, C. (2006). *Moral responsibility*. Oxford: Routledge.
- Moya, C. (2007). Moral responsibility without alternative possibilities?. *Journal of Philosophy*, 104(9), 475–486.
- Moya, C. (2010). ‘Blame, moral obligation and alternative possibilities’, research talk: Spanish Society for Analytic Philosophy (SEFA). Tenerife: Universidad de La Laguna.
- Moya, C. (2011). On the very idea of a robust alternative. In *Crítica, Revista Hispanoamericana Filosofía*, 41 (August): 3–26.
- Nelkin, D. (2011). *Making sense of freedom and responsibility*. Oxford: Oxford University Press.
- Otsuka, M. (1998). Incompatibilism and the avoidability of blame. *Ethics*, 108(4), 685–701.
- Owens, D. (2000). *Reason without freedom*. London: Routledge.
- Pereboom, D. (2000). Alternative Possibilities and Causal Histories. *Philosophical Perspectives* 14, papers on Action and Freedom, James E. Tomberlin (ed.): 119–37.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Pereboom, D. (2009). Further thoughts about a Frankfurt-style argument. *Philosophical Explorations*, 12(2), 109–118.
- Pereboom, D. (2011). Free will, skepticism and meaning in life. In R. Kane (Ed.), *Oxford handbook of free will* (2nd ed.). New York: Oxford University Press.
- Pereboom, D. (2012). Frankfurt examples, derivative responsibility, and the timing objection. *Philosophical Issues: A Supplement to Nous*, 22(1), 298–315.
- Smart, J. J. C. (1959). Sensations and brain processes. *Philosophical Review*, 68(2), 141–156.
- Watson, G. (2003) (ed.) *Free Will*, Oxford Readings in Philosophy 2nd edn. Oxford: Oxford University Press.
- Widerker, D. (1995). Libertarianism and Frankfurt’s attack on the principle of alternative possibilities. *Philosophical Review*, 104(2), 247–261.
- Widerker, D. (2000). Frankfurt’s attack on the principle of alternative possibilities: a further look. *Philosophical Perspectives* 14, papers on Action and Freedom, James E. Tomberlin (ed.): 181–201.
- Widerker, D. (2006). Libertarianism and the philosophical significance of Frankfurt scenarios. *Journal of Philosophy*, 103, 169–187.
- Widerker, D. (2009). A defense of Frankfurt-friendly libertarianism. *Philosophical Explorations*, 12(2), 87–108.
- Widerker, D. (2011). Frankfurt-style libertarianism. In R. Kane (Ed.), *Oxford handbook of free will* (2nd ed.). New York: Oxford University Press.
- Wolf, S. (1990). *Freedom within reason*. Oxford: Oxford University Press.