

Coercion and Volition in Law and Philosophy

Amit Pundik

Introduction

Emil Bednarek, a Polish political prisoner and senior block leader at Auschwitz, was convicted at the second Auschwitz Trial in 1965 of fourteen murders he had committed during the Holocaust.¹ Bednarek was found to have shown a “lust for killing” (*Mordlust*). In addition to zealously enforcing the rules of the camp, he went further and created forms of torture without orders from his superiors. In addition, he would drag the weaker prisoners (*Muselman*) into his block and beat them to death when they crossed his path, and play sadistic games with his victims, tormenting them both physically and psychologically. A striking feature of this case is that Bednarek was indicted *only* for those actions he was not ordered to perform—not for those required by his superiors.² Putting this scenario in terms of *causal overdetermination*, when a certain action is induced by both external coercion and the defendant’s own volition,³ the presence of the defendant’s own volition is ignored and the action is treated similarly to that induced solely by external coercion.

Like Bednarek, the rest of the defendants in the second Auschwitz Trial (all of whom were officials and guards) were also indicted only for actions they were not coerced to perform. While the second trials attracted fierce criticism,⁴ and the question of whether Nazi soldiers and officers committed the mass murders under coercion has been subject to an intense debate,⁵ the legal approach that external coercion renders the defendant’s volition immaterial has not been challenged or even acknowledged directly.

1. Frankfurter Auschwitz-Prozess, 19 August 1965, Strafsache gegen Mulka ua, 4 Ks 2/63 Landgericht Frankfurt am Main, 182, online: <http://www.auschwitz-prozess.de/index.php> [Bednarek] (Oral reasoning of the Chairman of the Judiciary, Hans Hofmeyer). Bednarek was accused of more than fourteen murders but he was not convicted on all counts due to lack of sufficient evidence. An accessible account of the case is available in Rebecca Wittmann, *Beyond Justice* (Harvard University Press, 2005) at 233-35.
2. Given that he was involved in the killing and assault of many more people, it is difficult to accept that evidence was found only of actions he performed of his own accord but *never* of those he was ordered to perform.
3. Note that the terms “volition” and “want,” which are common in the Philosophy of Action literature, do not introduce new requirements over and above the elements required by the Criminal *mens rea*. In particular, they do not require reference to the aims or motives of the action, which Criminal Law traditionally ignores.
4. Much criticism was levelled against the court’s decision to assess the defendants’ actions using the Criminal Law that was in effect when the actions were performed, namely the Nazi Criminal Code. Wittmann, for example, claims that this made “the prosecution dependent on the same standards of illegality the Nazis themselves had used to investigate criminal activity in the camps. This reliance on the letter of the law legitimated the criminal Nazi state and set a standard for illegal behavior in the 1960s Frankfurt courtroom that eerily echoed the laws of the Third Reich.” Wittmann, *supra* note 1 at 272.
5. See, most notably, the controversy surrounding Daniel J Goldhagen, *Hitler’s Willing Executioners: Ordinary Germans and the Holocaust* (Vintage Books, 1997).

A similar approach to coercion and volition is currently taken by various jurisdictions, as exemplified by their definitions of the criminal defence of duress, which do not limit the application of this defence to cases in which the defendant did not want to perform the criminal actions. External coercion renders the defendant's own volition immaterial, according to contemporary German Criminal law.⁶ In the United States, the key factor in applying duress is the quality of the coercive force or threat (that is, with what harm the defendant was threatened, and whether their inability to resist the coercion was reasonable).⁷ The more detailed Australian requirements focus on the quality of the coercion,⁸ while also including a causal connection between the coercion and the commission of the crime.⁹ A similar requirement of causal connection appears in English Criminal Law.¹⁰ Neither Australia nor England requires an absence of an additional causal link between the defendant's own volition and the criminal action. Therefore, it seems safe to conclude that, as a matter of positive law, cases in which the accused *also* wanted to commit the crime of their own volition are not excluded from the defence of duress.¹¹

An immediate and plausible explanation for Criminal Law's approach to duress is that criminal responsibility requires *alternative possibilities*, namely that the defendant had a choice between two or more (acceptable)¹² courses of action. Bednarek, being himself a prisoner at Auschwitz, clearly did not have much choice over whether to obey the orders he was given, and was thus absolved of criminal—and perhaps also moral—responsibility for the actions he was required to perform. As the grounds for his convictions indicate, these actions were probably not induced (only) by the lack of choice but (also) by his own volition (be it his “lust for killing” or something else). Nevertheless, the lack of alternatives was probably one of the prosecution's reasons for ignoring his volition (that he most likely *wanted* to do what he had no alternative but to do).

6. §34 and §35 Strafgesetzbuch (German Penal Code).

7. See, for example, section 2.09(1) of the Model Penal Code: “It is an affirmative defense that the actor engaged in the conduct charged to constitute an offense because he was coerced to do so by the use of, or a threat to use unlawful force against his person or the person of another, that a person of reasonable firmness in his situation would have been unable to resist.”

8. “Where the accused has been required to do the act charged against him (i) under a threat that death or grievous bodily harm will be inflicted unlawfully upon a human being if the accused fails to do the act and (ii) the circumstances were such that a person of ordinary firmness would have been likely to yield to the threat in the way the accused did and (iii) the threat was present and continuing, imminent and impending (as previously described) and (iv) the accused reasonably apprehended that the threat would be carried out and (v) he was induced thereby to commit the crime charged and (vi) that crime was not murder, nor any other crime so heinous as to be excepted from the doctrine and (vii) the accused did not, by fault on his part when free from the duress, expose himself to its application and (viii) he had no means, with safety to himself, of preventing the execution of the threat, then the accused, in such circumstances at least, has a defence of duress.” *R v Hurley & Murray*, [1967] VR 526 at 543.

9. See in particular the fifth requirement, *ibid*.

10. *R v Cole*, [1994] Crim LR 582.

11. One could argue that, even without an explicit exclusion, the defence of duress does not apply to such cases (for instance, because all excuses arguably require that every defendant would not have acted the way they did but for the excuse). Such a claim, however, finds no support in the legislation or the jurisprudence of the aforementioned jurisdictions.

12. The distinction between ‘lack of actual alternatives’ and ‘lack of acceptable or reasonable alternatives’ is discussed in the text accompanying note 60 below.

The lack of alternative possibilities plays a key role in the scholarly accounts of the defence of duress. This rationale has accompanied Common Law from its early days. Francis Bacon, for example, describes duress as being present in cases in which “either there has been an impossibility for a man to do otherwise, or so great a perturbation of the judgement as in presumption of law man’s nature cannot overcome.”¹³ It also appears in the accounts of modern legal scholars.¹⁴ H.L.A. Hart holds that, in cases of coercion, “it is morally wrong to punish because ‘he could not have helped it’ or ‘he could not have done otherwise’ or ‘he had no real choice’,”¹⁵ and Michael Moore assigns a central role to the notion of ‘could have done otherwise’ in his choice theory of excuse.¹⁶ While this view has not been unanimously accepted,¹⁷ both its proponents and opponents accept it as the orthodox view of duress.¹⁸ And while duress is a paradigmatic example for excusing the agent from criminal responsibility due to their lack of alternatives even when they wanted to do what they did, a similar logic arguably applies to the defence of necessity,¹⁹ and to any other case in which the agent is excused due to lack of alternatives.

In stark contrast to the legal approach to coercion and volition, philosophical discussion has been swayed in the opposite direction. In a ground-breaking paper, Harry Frankfurt challenged the Principle of Alternative Possibilities (PAP), according to which “a person is morally responsible for what he has done only

-
13. Francis Bacon, *The Elements of the Common Laws of England* (London, 1630) in *Works* vol 8, ed by Montague (London, 1831) at 131.
 14. Importantly, none of these accounts discusses the question of coinciding coercion and volition. An analogous question arises in the different context of self-defence, whenever A intends to harm B not in order to eliminate B’s attack on A but rather out of malice. See the exchange between George Fletcher, *Rethinking Criminal Law* (Little, Brown, 1978) at 559-62 and Paul H Robinson, *Structure and Function in Criminal Law* (Clarendon Press, 1997) at 102, 108-11; see also Lawrence Crocker, “Justification and Bad Motives” (2008) 6:1 *Ohio State J Crim Law* 277. Note, however, that whatever the rationale of self-defence is, it is not rooted in the lack of alternative possibilities and thus this discussion is not relevant to this paper.
 15. “What is crucial is that those whom we punish should have had, when they acted, the normal capacities, physical and mental, for doing what the law requires and abstaining from what it forbids, and a fair opportunity to exercise these capacities. Where these capacities and opportunities are absent, as they are in different ways in the varied cases of accident, mistake, paralysis, reflex action, coercion, insanity, etc., the moral protest is that it is morally wrong to punish because ‘he could not have helped it’ or ‘he could not have done otherwise’ or ‘he had no real choice’.” HLA Hart, *Punishment and Responsibility: Essays in the Philosophy of Law*, 2nd ed (Oxford University Press, 2008) at 152. Note that Hart devoted the second chapter of his book to explaining why his approach to alternative possibilities does *not* require commitment to indeterminism.
 16. See Michael S Moore, *Placing Blame: A Theory of Criminal Law* (Oxford University Press, 2010) at ch 13.
 17. See, most notably, Victor Tadros, *Criminal Responsibility* (Oxford University Press, 2005) at 57-61 (and see note 45 below for the problem with Tadros’ key example).
 18. “The choice theory just examined has come to be regarded as the traditional or orthodox view of excuse.” Moore, *supra* note 16 at 562. “The central way in which capacity is commonly considered relevant has to do with alternative possibilities.” Tadros, *ibid* at 57.
 19. “[T]he similarities between the two defences [duress and necessity—AP] are so great that consistency and logic require that they be understood as based on the same juristic principles. Indeed, to do otherwise would be to promote incoherence and anomaly in the criminal law.” Lamer, CJ in *R v Hibbert* (1995) 2 SCR 973 at para 54 cited with approval in *R v Ryan* (2013) 1 SCR 14 at para 17. I am grateful to the anonymous referee for drawing my attention to this similarity.

if he could have done otherwise,”²⁰ which was taken for granted at the time. Through a series of intuitive counterexamples involving coercion, he aims to demonstrate that a person can be morally responsible even if coerced (and hence has no alternative possibilities), and thus the PAP is not a condition of moral responsibility. Frankfurt seeks to establish that the crucial factor in assessing the individual’s responsibility for their actions is not the alternatives they have or lack, but whether their action was motivated by their own volition. So long as the agent acted on their own volition, they may be responsible, regardless of whether they could have done otherwise. According to Frankfurt, Bednarek is morally responsible for any action he performed willingly, even for those actions which he had no alternative but to perform. Frankfurt’s attack on the PAP challenges not only the outcome in cases like that of Bednarek, but also the significance given to the availability of alternative possibilities by Hart and Moore.

This paper seeks to vindicate the legal approach to coercion and volition by showing that the role of alternative possibilities in determining the defendant’s responsibility differs from that assigned to them by both Hart and Moore, on the one hand, and by Frankfurt and his commentators, on the other. It is suggested here that the principle that underlies the legal approach is not the PAP, but rather a more nuanced version of it, according to which a person is *not* responsible for what they have done if they (justifiably)²¹ *believed* they could *not* have done otherwise.²² While Bednarek might be responsible for choosing the wrong reason for action or having a positive disposition toward the coerced action, he is *not* responsible for performing the action itself. According to this principle, termed here the Epistemic Principle of Alternative Possibilities (E-PAP), the responsibility of persons is determined not by their ontological position (the alternative possibilities they actually have) but purely by their *epistemic* position (their beliefs about their alternative possibilities).²³ This paper argues that the E-PAP is immune to Frankfurt’s counterexamples. Hence, even if, in theory, one can be responsible for one’s actions even when one has no alternative, as Frankfurt claims (and I take no stance on the matter), Frankfurt’s attack does not pose

-
20. Harry Frankfurt, “Alternate Possibilities and Moral Responsibility” (1969) 66:23 *Journal of Philosophy* 829 at 829.
 21. The term “justifiably” is bracketed because jurisdictions vary in their approach to the required belief, namely whether it must be merely sincere or also reasonable. See the text accompanying note 54 below.
 22. An important qualification to this phrasing is that a person who killed another person to save one’s own life may still be criminally responsible for murder (as in the famous case of *R v Dudley and Stephens* (1884) 14 QB 273 (in which two sailors whose ship sank killed and ate a 17-year-old boy after 20 days lost at sea). Identifying the rationale underlying such an exception lies beyond the scope of this paper.
 23. Cf Kimberly Kessler Ferzan, “Justifying Self-Defense” (2005) 24:6 *Law & Phil* 711. Ferzan argues that self-defence contains objective conditions and subjective limitations, thereby locating “the appropriate epistemic vantage point in the self-defender,” and concluding that “an inquiry into the reasonableness of the defender’s belief is not required” (*ibid* at 714-15). See also Helen Frowe, “A Practical Account of Self-Defence” (2010) 29:3 *Law & Phil* 245. Like Ferzan, this paper seeks to highlight the significance of epistemic considerations. Note, however, that the analogy between the defences of duress and self-defence is problematic, partly because arguably the former is an excuse while the latter a justification, and partly because a coerced person who also wants to commit the crime is *not* “...doing the right deed for the wrong reason.” Fletcher, *supra* note 14 at 556.

a threat to the legal approach's focus on the external coercion rather than on the internal volition. The significance of the issue of coercion and volition goes beyond cases like *Bednarek*: by focusing on the agent's subjective beliefs, the E-PAP highlights the leading role that moral and criminal responsibility ascribe to the agent's internal point of view, even when the external reality differs from their beliefs.

The first section of this paper seeks to show that the force of Frankfurt's attack on the PAP stems exclusively from one example alone. This is because his other examples, including the one most similar to *Bednarek*, pose no serious challenge either to the traditional PAP or to the legal approach. The second section focuses on Frankfurt's strongest example and argues that he and his commentators have neglected an important feature that appears in most legal cases of coercion: the fact that the agent *believed* they had no alternative. It is shown that once the epistemic element of belief is integrated into the role of alternative possibilities, Frankfurt's seemingly-strongest example becomes ineffective. The third section explores the E-PAP in more detail and illustrates that existing legal doctrines reflect the E-PAP rather than the PAP. The last section discusses two potential objections to this paper's attempt to bring together the philosophical and legal discussions.

Between Jones and Bednarek

Frankfurt sought to dispel what was then a commonly held view, according to which the absence of alternative possibilities automatically absolves the agent of moral responsibility. To do so, he described a series of intuitive counterexamples that aimed to demonstrate that a person can be morally responsible even if they have no alternative, and hence the PAP is not a condition of moral responsibility. The best example to begin with is that of Jones₄, which is Frankfurt's strongest. In this example, Black wants Jones₄ to perform a certain action. If it becomes clear to Black that Jones₄ is going to decide to do something different, Black will make sure Jones₄ decides to, and then does, perform the action. However, Black prefers to avoid interfering unnecessarily, so he will do nothing if Jones₄ performs this action of his own accord. Jones₄ would thus become aware of Black's interference only once Black actually interfered.

Frankfurt persuasively argues that if "Jones₄, for reasons of his own, decides to perform and does perform the very action Black wants him to perform [...] Jones₄ will bear precisely the same moral responsibility for what he does as he would have borne if Black had not been ready to take steps to ensure that he do it."²⁴ Jones₄ is an effective counterexample to the traditional PAP because, given that "[w]hatever Jones₄'s initial preferences and inclinations, [...] Black will have his way,"²⁵ the PAP counterintuitively implies that Jones₄ is not morally responsible. This forceful counterexample has led Frankfurt to conclude that the PAP cannot be right.

24. Frankfurt, *supra* note 20 at 836.

25. *Ibid* at 835.

Frankfurt's attack on the PAP generated considerable attention within the philosophical literature.²⁶ It led to various objections²⁷ and numerous attempts to devise variations on Jones₄ that would be immune to these objections,²⁸ setting in motion yet another set of objections and replies.²⁹ The important point to note is that Frankfurt's attack on the PAP has led philosophers to develop new theories of moral responsibility that explain how agents can be responsible for their actions even when they have no alternative. Frankfurt himself, for example, later proposed that a person is morally responsible only if they had a (second-order) desire to have the (first-order) desire to act as they did.³⁰ The agent's second-order desire can be aligned with their first-order even when they have no alternative but to act as they do. Another example is Fischer and Ravizza's influential theory that distinguishes between "regulative control," which does require alternative possibilities (but is not required for moral responsibility), and "guidance control," which suffices for moral responsibility and is based on the agent's responsiveness to reasons (and an agent is arguably able to respond to reasons even in the absence of an alternative).³¹

But Emil Bednarek was not like Jones₄. In the case of Jones₄ it is clear that his decision was induced exclusively by his own volition, and the coercive back-up mechanism that was in place had no influence on this decision. The causal pathway in Jones₄, namely, how the action was *actually* induced, does not involve the coercive mechanism. By contrast, in Bednarek's case, the actual causal pathway that led to his actions is less clear, since both the coercive orders and his own volition were present in the situation and it is not easy to determine which induced his actions. On the one hand, his superiors' orders left him with no alternative but to obey,³² but, on the other hand, he also wanted to act as he did. *Bednarek* is a challenging case because it combines two intuitive ingredients of moral and

-
26. An extensive body of literature has developed around Frankfurt's examples. See, for example, the collection of papers in David Widerker & Michael McKenna, eds, *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (Ashgate, 2003). The following are but a few examples.
 27. For just two examples, see David Widerker, "Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities" (1995) 104:2 *Philosophical Rev* 247 and Carl Ginet, "In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Argument Convincing" (1996) 10 *Philosophical Perspectives* 403.
 28. See, for example, Alfred Mele & David Robb, "Rescuing Frankfurt-Style Cases" (1998) 107:1 *Philosophical Rev* 97; Derk Pereboom, "Alternative Possibilities and Causal Histories" (2000) 14 *Philosophical Perspectives* 119.
 29. See David Widerker, "Frankfurt's Attack on Alternative Possibilities: A Further Look" (2000) 14 *Philosophical Perspectives* 181; Alfred R Mele & David Robb, "Bbs, Magnets and Seesaws: The Metaphysics of Frankfurt-style Cases" in David Widerker & Michael McKenna, eds, *Moral Responsibility and Alternative Possibilities* (Ashgate, 2003) 127. While the intense discussion that ensued from Frankfurt's paper has raised various important issues, this paper focuses on the original Jones₄ rather than on any of the other, more complicated (and often more surreal) hypothetical examples, because they are all similar to Jones, with respect to the feature central to this paper (the agent's knowledge of their lack of alternative).
 30. See "Freedom of the Will and the Concept of a Person" (1971) 68:1 *Journal of Philosophy* 5. For a critical discussion, see, for example, Gary Watson, "Free Agency" (1975) 72:8 *Journal of Philosophy* 205.
 31. John M Fisher & Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge University Press, 1998) at 31-34.
 32. For the claim that he could have done otherwise, namely disobey and suffer the lethal consequences, see the text following note 41 below.

criminal responsibility,³³ each pushing in the opposite direction: acting willingly (which is present in *Bednarek*) and choosing between alternative courses of action (which is not).

Frankfurt begins his paper with three hypothetical examples that explore situations in which these two ingredients are in conflict: the agent is exposed to a coercive mechanism that leaves them with no alternative but to do what they wanted to do anyway.³⁴ The first two are extreme examples in which one of the ingredients is muted. Jones₁ is an unreasonable man, “who does what he has once decided to do no matter what happens next and no matter what the cost.”³⁵ Jones₁ is clearly responsible for acting the way he did even though he was exposed to the coercive mechanism. However, Frankfurt rightly points out that this case is not a counterexample to the traditional PAP, for the reason that “the threat had no coercive effect and, hence, that it did not actually deprive him of alternatives to doing what he did.”³⁶ Indeed, as a matter of positive law, if Jones₁ were put on trial for his actions, the defence of duress would not be available to him in any jurisdiction that required a causal connection between the coercion and the commission of the crime (such as Australia or England). The lack of alternative does not prevent the conclusion that Jones₁ is responsible, because even if Jones₁ lacked an alternative when doing what he did, it was only because of his prior decision to act that way. A person who is responsible for positioning themselves in a situation in which they have no alternative but to act in the way they previously decided is responsible for acting on their previous decision. So even though the two intuitive ingredients of responsibility are in conflict in the case of Jones₁, this case does not pose a serious challenge to either the PAP or the defence of duress because the lack of choice is muted by Jones₁'s earlier decision.

In Frankfurt's second example, it is the volitional ingredient that is muted. Jones₂ “was stampeded by the threat. Given that threat, he would have performed that action regardless of what decision he had already made.”³⁷ Frankfurt concedes that Jones₂ “can hardly be said to be morally responsible for his action,”³⁸ and thus cannot serve as a counterexample to the PAP because the threat denied him of any alternative but to obey. As for positive law, Jones₂ seems like a clear case of acting under duress. While he had already formed the intention to perform the action, when the time to act arrived, his volition was stampeded by the coercive mechanism. Jones₂ does not pose a serious challenge to either the PAP or the defence of duress because the other ingredient of responsibility is muted: Jones₂ did not act willingly.

Frankfurt's third example is closest to the case of *Bednarek*. Jones₃ “was neither stampeded by the threat nor indifferent to it. [...] When he acted, he was not actually motivated by the threat but solely by the considerations that had

33. For the relation between moral and criminal responsibility, see the text accompanying note 60 below.

34. Frankfurt *supra* note 20 at 830-33.

35. *Ibid* at 831.

36. *Ibid* at 832.

37. *Ibid*.

38. *Ibid*.

originally commended the action to him.”³⁹ Jones₃ is a cleaner example than that of Bednarek, yet one that is also oversimplified, because its causal pathway is stipulated: Jones₃’s actions were induced by his own volition, as he acts *solely* on his original considerations. By contrast, it is not fully clear whether Bednarek’s actions were induced by the external coercion, his own volition or both. After all, real people are sometimes motivated by more than one reason simultaneously, and even if either the external coercion or the individual’s own volition suffices to motivate the action on its own, it does not mean they cannot operate together (thereby leading to a causal overdetermination). Yet Frankfurt’s Jones₃ is a good example to explore, because it crystallises the tension between the two ingredients. Unlike the previous examples, neither ingredient is muted—and yet the causal pathway is kept clear and simple. Not only did Jones₃’s volition take part in the causal pathway of his actions (as it surely did in *Bednarek* too); it also had an exclusive role in inducing Jones₃’s actions, while the coercive mechanism played no role at all.⁴⁰ Jones₃ is clearly an unrealistic example because a person who is exposed to an effective coercive mechanism is unlikely to act without giving *any* consideration to the coercive force or threat to which they are subjected. However, Jones₃ provides an instructive test case for the legal approach to coercion and volition, which ignores the defendant’s volition whenever their actions were coerced. If the defence of duress applied even to Jones₃, it would apply to any other case in which the two ingredients conflict, because the contribution of the defendant’s volition to their actions could not be any greater than in the case of Jones₃.

According to Frankfurt, Jones₃ seems to provide a “decisive counter-example” to the PAP: the threat denies Jones₃ the alternative of not performing the action, “yet the threat, since Jones₃ performs the action without regard to it, does not reduce his moral responsibility for what he does.”⁴¹ It seems that Frankfurt has the intuition (which I do not share) that Jones₃ *is* responsible for his actions, so Jones₃ would serve as a counterexample to the PAP if the latter were to lead to the conclusion that Jones₃ is not responsible.

But Frankfurt notes an immediate objection that could be levelled against Jones₃’s being a counterexample to the PAP. Jones₃ *can* do otherwise: he can defy the threat and accept the penalty Black will impose on him. As a result, Jones₃ is not a counterexample to the PAP because he could be responsible even according to the PAP: he had the (costly) alternative of defying the threat and hence he could be responsible for not choosing this alternative. Jones₃ is not a counterexample to the PAP even for those who share Frankfurt’s intuition about this case, because the PAP does not necessitate the conclusion that Jones₃ is not morally responsible.

39. *Ibid* at 832.

40. Notably, in Jones₁ too, the agent’s volition dominates the causal pathway. However, in Jones₁ the decision to act upon his volition is made prior to the coercive threat, and thus the latter *cannot* play a role in the causal pathway. By contrast, in Jones₃, in which the decision is made after the threat is issued, the threat can play a causal role and would have influenced Jones₃’s actions, had his volition been different.

41. *Ibid* at 834.

It is not surprising that Frankfurt seeks to bypass discussion of this objection in detail, because engaging in this discussion would show that Jones₃ is far from being the ‘decisive’ counterexample Frankfurt claims it to be. To see why, consider what such a discussion would look like. Recall that, to refute the objection, Frankfurt would need to show that the PAP ought to lead to the conclusion that Jones₃ cannot be responsible. As Frankfurt acknowledges, to show that the PAP yields this conclusion, it is necessary to delve into the analysis of the concept of “could have done otherwise” and show that Black’s threat prevents Jones₃ from doing otherwise.⁴² The concept of “could have done otherwise” has been subject to thorough and technical discussion in the literature.⁴³ Instead of joining the crowded discussion of this analysis, Frankfurt proceeds straight to Jones₄, which he takes to be such a forceful counterexample to the PAP that it shows the latter to be false, no matter how “could have done otherwise” should be understood.⁴⁴ Frankfurt’s attack on the PAP thus hinges on the success of Jones₄. If the example of Jones₄ could be set aside—and the next section suggests how this indeed could be done—Frankfurt’s attack on the PAP would fail to offer any argument as to why either Jones₃ or Bednarek should not be absolved of responsibility on the grounds they could not have done otherwise.

Furthermore, Frankfurt’s supposedly intuitive starting point that Jones₃ is responsible is questionable, because this intuition can be explained away through a careful examination of what exactly Jones₃ is responsible *for*. The coercive threat did not make Jones₃ form the intention to perform the action prior to learning about the threat. Nor did it make Jones₃ choose his own desire as his sole or main reason for action or form a positive mental attitude toward his action while carrying it out. It could be argued that, while it seems intuitive that Jones₃ is responsible for performing the action, what he is really responsible for is either forming the intention or choosing the wrong reason for action. But if Jones₃ is responsible only for these, then the intuition that he is responsible does not challenge the PAP because the PAP does not yield counterintuitive conclusions with regard to forming intentions and choosing reasons for action. Jones₃ had alternative possibilities with regard to these, and thus can be responsible for them. What the coercive threat left Jones₃ with no alternative but to do was to perform the action. Therefore, for Jones₃ to serve as a successful counterexample to the PAP, it has to be shown that he is responsible *for performing the action* rather than for anything else. But once the question is refocused on the action itself, it is no longer intuitive that Jones₃ could be responsible for something that he could not avoid doing.⁴⁵

42. *Ibid* at 834–35.

43. One common strategy, known as the Conditional Analysis, is to hold that a person could have done otherwise if they would have done otherwise *had they wanted to*. RE Hobart, “Free Will as Involving Determination and Inconceivable without It” (1934) 43:169 *Mind* 1. For a forceful objection, see Peter van Inwagen, *An Essay on Free Will* (Oxford University Press, 1983) at 114–26.

44. Frankfurt, *supra* note 20 at 835.

45. Interestingly, Victor Tadros raises a similar example to Jones₄, which is also meant to show that alternative possibilities are not needed for (criminal) responsibility, and Tadros’ discussion is more nuanced with regard to what the agent is responsible *for*. In Tadros’ example,

This section has shown that none of first three Joneses poses any serious challenge to the legal approach to alternative possibilities, according to which lack of alternative suffices to render the agent not responsible, even if he acted willingly. Jones₃, which is the example most similar to *Bednarek*, is much less of a ‘decisive’ counterexample to the PAP than Frankfurt holds it to be: rather than providing an argument, it is based on sharing a certain intuition that could be explained away. If the current philosophical approach (which has been swayed by Frankfurt’s attack on the PAP) is to be preferred to the legal approach, it is only because of Jones₄, to which the next section turns.

Hidden Coercion: The Case of Jones₄

Recall that in Jones₄, there is a coercive mechanism in place that will make him perform a certain action, but this mechanism will be triggered only if he does not perform this action of his own accord. Frankfurt argues that, if Jones₄ decides to perform the action of his own accord, he will be morally responsible, and hence Jones₄ serves as an effective counterexample to the PAP, which yields the counterintuitive conclusion that Jones₄ is not morally responsible (because he could not have done otherwise).

Jones₄ is not only unlike *Bednarek*; it is also unlike most real cases of duress. Even though the coercive mechanism deprived Jones₄ of any alternative possibilities, Jones₄ was unaware of that fact. Jones₄ did not *believe* he could not have done otherwise, because when he performed the action he did not believe, nor did he have any reason to suspect, that Black would step in if Jones₄ decided to act differently. By contrast, *Bednarek* knew all too well the consequences of disobeying his Nazi superiors in Auschwitz. Furthermore, in many jurisdictions, no real case in which the defendant was exposed to a coercive mechanism of which they were unaware has ever been litigated.⁴⁶ Frankfurt’s example contains a feature that could not be found in legal cases in the present investigation, and is unlikely to be found in real-life praxis.

Had the feature of lack of belief been removed from Jones₄, this example would have become similar to Jones₃ and *Bednarek*. The conclusion the PAP

Derek is locked in a room with a single door that he believes to be open. Through the small window, he sees his child crawling toward a pool of water but, being an uncaring parent, he lets the child go on without trying to leave the room to stop them, and the child drowns. Tadros concedes that “the fact he could not leave the room might be thought to break the chain of causation between his staying in the room and the child drowning.” Tadros, *supra* note 17 at 63. However, Tadros insists that “he is at least responsible for staying in the room” (*ibid*). While Tadros focuses on what Derek is responsible for, it is unclear what the basis for his conclusion is. Sure, Derek is responsible for *not trying* to leave the room. However, Tadros provides no argument as to why Derek is responsible for *staying in the room* if he could not have left it. On the contrary, his concession about the possible break in the chain of causation is equally applicable here: if the locked door broke the connection between his staying in the room and the drowning of the child, it also broke the connection between his not trying to leave and his staying in the room.

46. Searches were conducted in the following jurisdictions: Australia, Canada, England and Wales, Germany, Israel and the United States (though only of judgements referring to the Model Penal Code).

would have yielded if Jones₄ had believed that he was exposed to coercion is no longer problematic, or at least Frankfurt provides no reason to think it is. The feature of Jones₄'s lack of belief is thus constitutive of Frankfurt's attack on the PAP.

But if this feature does not appear in real cases, it is possible to identify an alternative principle that would capture the realistic cases of coercion: that a person is *not* responsible for what they have done if they (justifiably) *believed* they could *not* have done otherwise (that is, the aforementioned E-PAP). The E-PAP yields that the defendant's responsibility is negated not when they actually lacked an alternative but when they (justifiably) so believed. Like the PAP, the E-PAP captures the important role that alternative possibilities play in our current legal practices; but, unlike the PAP, it does not apply to unrealistic cases like Jones₄. Maybe Frankfurt is right about Jones₄ and maybe he is not. Either way, accounting for the role of alternative possibilities in these hypothetical cases is a different project from that involved in addressing real cases like Bednarek.

Cases like Jones₄ are not counterexamples to the E-PAP because the latter does not yield the counterintuitive implication that Jones₄ is not morally responsible. An important feature of this example is that Jones₄ is unaware of the fact that he is left with no alternative but to perform the action. This is where the traditional PAP and the E-PAP diverge: while the PAP makes responsibility hinge on the absence of alternative possibilities, the E-PAP requires the agent's belief in that absence. The E-PAP is thus a principle that incorporates the lack of alternative possibilities in judgements of responsibility, but it does so in a way that does not counterintuitively imply that Jones₄ is not morally responsible.

When the E-PAP is applied to Jones₄, it becomes clear that part of Jones₄'s intuitive force is derived from Frankfurt's focus on acting, while neglecting the importance of choosing. An important difference between Frankfurt's analysis of Jones₄ and how the E-PAP applies to this example lies in the sensitivity to what Jones₄ chooses to do. According to Frankfurt's analysis, Jones₄'s responsibility is unaffected by his choice, because he will be coerced to perform the action irrespective of what he chooses to do. As a result, rather than being sensitive to Jones₄'s choice, Jones₄'s responsibility is sensitive to Black's choice regarding whether to impose these constraints on Jones₄. By contrast, the E-PAP emphasises the agent's choice. The agent's belief about the alternatives they do, or do not, have is what enables their choice to begin with: an agent who believes they have no alternative but to perform a certain action is not only in a position in which they "could not have done otherwise"; they are also in a position in which they cannot even engage in the process of choosing.⁴⁷ The E-PAP's focus on the

47. Immanuel Kant, for example, seems to hold that the process of deliberation requires the presupposition that the future is open. See MJ Gregor & CM Korsgaard, eds, *Immanuel Kant: Groundwork of the Metaphysics of Morals* (Cambridge University Press, 1998) at 54. The Kantian position can be understood as a psychological claim, according to which human beings are unable to deliberate while they believe there is no alternative course of action open to them. Notably, this claim seems to require the incorporation of epistemic elements into the principle governing alternative possibilities, because this explanation, too, focuses on the agent's epistemic position *vis-à-vis* their alternative possibilities.

agent's belief thus makes the analysis of Jones₄'s responsibility highly sensitive to Jones₄'s choice. If Jones₄ chooses to perform the action, then he is responsible, as shown here. However, if Jones₄ does not choose to perform the action, then although Black will eventually make him perform the action, the former will not be responsible for its performance.

Like Jones₄, the first three Joneses cannot serve as counterexamples to the E-PAP either. Jones₁, the unreasonable man “who does what he has once decided to do no matter what happens next and no matter what the cost,”⁴⁸ is certainly not a counterexample to the E-PAP. Frankfurt argues that the threat did not deprive Jones₁ of the alternatives he had. If this is true, nor did it deprive him of any belief he might have had with regard to these alternatives. Nor is Jones₂ (who “would have performed that action regardless of what decision he had already made”)⁴⁹ a counterexample to the E-PAP. Jones₂ is not responsible for what he has done because he *believed* he could not have done otherwise. Jones₃, who “was not actually motivated by the threat but solely by the considerations that had originally commended the action to him,”⁵⁰ could be viewed as a counterexample not only to the PAP but also to the E-PAP: Jones₃ was aware of Black's threat and hence could believe that he could not do otherwise, yet he is allegedly responsible for performing the action. However, Frankfurt raises no argument as to why Jones₃ is responsible for actions he could not have avoided, so Jones₃ is no more a counterexample to the E-PAP than it is to the PAP. The E-PAP is therefore not vulnerable to Frankfurt's attack.

The E-PAP is all that is needed to vindicate the legal approach to real-life cases like *Bednarek*. According to the E-PAP, Bednarek is not responsible for the actions he was ordered to perform because he (justifiably) *believed* he had no choice but to do as he was told. However, when he went further than he was ordered, the defence of duress is no longer applicable because he had a choice (namely to do *only* as he was told), and he *knew* it. The E-PAP would yield a similar conclusion about any case in which a person did something they believed they had no choice but to do, even if it turned out they also wanted to do it. Since the legal approach does not deal with cases like Jones₄, the E-PAP suffices to ground it.

Exploring the E-PAP

Having shown that the E-PAP, which is not vulnerable to Frankfurt's counterexamples, suffices to ground the legal approach, the purpose of this section is to analyse the E-PAP in greater detail and suggest how it underlies several elements in the defence of duress. It is important first to emphasise that the E-PAP incorporates epistemic considerations—namely, whether the person (justifiably) *believed* they could not have done otherwise. This is perhaps the most significant way in which the E-PAP diverges from the existing literature on moral responsibility

48. Frankfurt, *supra* note 20 at 831.

49. *Ibid* at 832.

50. *Ibid*.

and alternative possibilities.⁵¹ According to the E-PAP, the person's *moral* status (whether they are responsible) is affected not by their ontological position (the alternative possibilities they actually have) but purely by their *epistemic* position (their belief about their alternative possibilities).

A few clarifications might be helpful. Firstly, while the traditional PAP is formulated as a necessary condition of responsibility, the E-PAP is formulated as a responsibility-negating principle: a person is *not* responsible if they (justifiably) believed that they could not have done otherwise. This difference highlights that the E-PAP identifies *an exception* to responsibility. The negative formulation better suits the defence of duress, because a defence is an exception that absolves the agent of the responsibility they would otherwise bear. Secondly, just as the PAP is not the sole condition of responsibility, the E-PAP is not the sole exception to responsibility. As a result, a person might still be absolved of responsibility even if it is not true that they believed they could not have done otherwise (for example, if they lack capacity due to insanity). Lastly, just as the PAP examines whether a person has alternative possibilities *when choosing how to act*, the E-PAP exempts them only if they believed at the time they made their choice that they had no alternative. Discovering *after* making the choice that they had no alternative does not absolve them of responsibility.

Existing legal doctrines indicate that the principle reflected in the legal approach is more similar to the E-PAP than to the traditional PAP. While the PAP focuses on the external reality (the existence or absence of actual alternatives), the legal approach gives a primary role to the agent's internal perspective in assessing their responsibility. While Frankfurt and his commentators focus exclusively on the external reality when evaluating the agent's responsibility in the coercive examples they devise, the legal approach acknowledges that there are situations in which the other conditions alone suffice to substantiate the defence of duress. To see that, consider the following example. Suppose that Black threatened Jones_s with a gun, which, unbeknownst to Jones_s, had a hidden malfunction. Even if Jones_s could have ignored Black's threat, Jones_s believed he had no alternative but to do as Black demanded. While, according to the PAP, Jones_s *may be* responsible because he could have done otherwise,⁵² according to

-
51. Notably, epistemic conditions of moral responsibility have previously appeared in the literature on the PAP. McKenna and Widerker introduce an epistemic condition of moral responsibility, according to which the agent "must have had some understanding of (or at least she must have been *able* to understand) the moral significance of her behaviour" (*supra* note 26 at 2). Similarly, Pereboom introduces a condition of robustness, according to which the agent "could have willed something different from what she actually willed such that she understood that by willing it she would thereby be precluded from moral responsibility for the action." Derk Pereboom, *Living Without Free Will* (Cambridge University Press, 2001) at 26. Both of these conditions assume that the agent knows they have an alternative, and focus on the agent's understanding of the moral implications of choosing that alternative. By contrast, the E-PAP focuses on the agent's knowledge of the very existence of an alternative.
52. It could be argued that Jones_s still lacked an alternative because of the psychological pressure created by the threat. This move would bring the PAP much closer to the E-PAP by injecting an epistemic element into the PAP because in such a case the agent's psychological pressure is strongly connected to their belief that the coercion is genuine and that they thus lack an alternative.

the E-PAP, Jones₅ is *not* responsible for doing what Black asked. This is because he (justifiably) believed he could not have done otherwise, *even if, in fact, he could have*. Had Jones₅ been prosecuted, he would have been excused on the grounds of the doctrine of ‘mistake of fact’,⁵³ which applies, in this case, to one of the facts establishing the defence of duress. The E-PAP is thus more suitable to the legal approach, which acknowledges that the defendant’s responsibility does not depend on the external facts but rather on their beliefs (mistaken or not) about these facts.

The focus on the defendant’s belief raises the immediate question of whether any sincere belief in the lack of an alternative should suffice to constitute the defence of duress, or whether this belief should satisfy certain conditions (for example, reasonableness). Assume that Jones₆ believed that Black would coerce him to perform a certain action he did not want to perform, yet he formed this belief because he is paranoid or simply used unreliable evidence about Black. In fact, Black was not really interested in what Jones₆ would do. Convinced by his belief, Jones₆ went on to perform the action. Should the defence of duress be available to Jones₆? While the legal approach grapples with this question, the PAP turns a blind eye to it.⁵⁴ According to the PAP, Jones₆ may be responsible simply because he clearly could have done otherwise.⁵⁵ That the E-PAP focuses on the defendant’s belief, thereby raising the question of its reasonableness—ignored by the PAP—illustrates yet again that the E-PAP is more suitable for grounding the legal approach.

Between the Criminal and the Moral

One might argue that there is little point in considering the philosophical scholarship on Frankfurt’s examples when discussing the legal approach to coercion and volition, because criminal and moral responsibility are such distinct phenomena that the philosophical and legal discussions are bound to diverge starkly. However, it is difficult to accept that convicting someone like Bednarek of murder does not require, reflect, imply or claim that Bednarek is morally responsible for the murder. Tadros holds that holding someone criminally responsible is just a specific instance of the general moral practice of holding agents responsible for what they do.⁵⁶ Husak describes the Anglo-American tradition as one that “borrows heavily from moral philosophy to illustrate the central issues in criminal theory,” and states that “[t]he legal analyses of responsibility, desert, wrongdoing, justification, and excuse are closely related to their counterparts in the

53. Andrew Ashworth, *Principles of Criminal Law*, 6th ed (Oxford University Press, 2009) at 215-19; Andrew P Simester et al, *Simester and Sullivan’s Criminal Law: Theory and Doctrine*, 6th ed (Bloomsbury, 2016) at 696-705.

54. The issue of false justified beliefs is contentious also philosophically, but most discussions appear in Epistemology; see, for example, Richard Feldman, *Epistemology* (Pearson, 2003) at ch 2; Clayton Littlejohn, *Justification and the Truth-Connection* (Cambridge University Press, 2012).

55. The point made in note 52 is also applicable here.

56. Tadros, *supra* note 17 at 23.

domain of morality.”⁵⁷ Moreover, this connection is particularly strong in the context of excuses. As Morse highlights, even if there are differences between criminal and moral responsibility, “the doctrines that excuse or mitigate criminal responsibility ... closely track the variables commonly thought to create moral excuse or mitigation.”⁵⁸ Evidence for this strong connection could also be found in the discussions of Hart and Moore of criminal excuses, both of which include detailed consideration (and rejection) of the threat determinism poses to criminal responsibility.⁵⁹ Whatever the precise connection between criminal and moral responsibility is, the legal approach to coercion and volition cannot afford to ignore the discussion of the moral responsibility of the agent in such cases.

A more specific objection to the relevance of the philosophical discussion to the legal approach to coercion and volition would point out that Frankfurt and his commentators are interested in the lack of *actual* alternatives, while the law is occupied mainly with the lack of *acceptable* alternatives.⁶⁰ For example, when a bank clerk is threatened with having their knees broken if they do not provide the code to the safe to a criminal who plans to burgle the bank, the former does have an alternative to providing the code, namely having their knees broken. However, if they choose to reveal the code, Criminal Law does not hold them responsible for doing so, because it recognises that it is unreasonable to expect them to choose the alternative. Thus, instead of asking whether the agent could have done otherwise, Criminal Law asks whether it was *reasonable to expect* them to have done otherwise. If the legal discussion is concerned with the reasonableness of the alternatives, while the philosophical discussion is interested in the actuality of the alternatives, it is unsurprising the two approaches select different examples, examine different principles and reach different conclusions.

However, whatever Frankfurt’s original examples referred to, his examples could easily be converted to cases of lack of acceptable alternatives, in which Jones has alternative possibilities but none of them is acceptable. Such converted examples would pose a similar challenge to the legal approach to the lack of acceptable alternatives as Frankfurt’s original examples allegedly posed to the

-
57. Douglas Husak, Book Review of *Punishment and Freedom: A Liberal Theory of Penal Justice* by Alan Brudner, (2010) 120:4 *Ethics* at 841 at 842. Husak believes that “political philosophy must be used in addition to moral philosophy to explain and justify the shape of Anglo-American law” (*ibid* at 844). However, Husak remains critical of Brudner’s Hegelian theory of punishment, according to which “culpability depends exclusively on what external choice a person has made, [so] everything pertaining to his inward self is irrelevant to a judgment of culpability.” Alan Brudner, *Punishment and Freedom: A Liberal Theory of Penal Justice* (Oxford University Press, 2009) at 166. See also David Shoemaker’s less unequivocal view in “On Criminal and Moral Responsibility” in Mark Timmons, ed, *Oxford Studies in Normative Ethics*, 3rd vol (Oxford University Press, 2013).
58. Stephen J Morse, “Psychopathy and Criminal Responsibility” (2008) 1 *Neuroethics* 205 at 208.
59. Hart, *supra* note 15 at ch 2; Moore *supra* note 16 at ch 12.
60. “In the standard case of coercion, B does X because B rationally regards X as the most attractive alternative—under the circumstances.” Alan Wertheimer, *Coercion* (Princeton University Press, 1987) at 10. Note that the controversy about what makes alternatives unacceptable or unreasonable is immaterial to this paper. In particular, it is immaterial whether the unacceptability or unreasonableness of the alternatives ought to be assessed empirically or normatively (see the extensive discussion of coercion as a moralised concept by Wertheimer).

philosophical approach to the lack of actual alternatives. If the clerk provides the code to the safe to a criminal who plans to burgle the bank, and does so for their own reasons (they hate the bank, they hope to receive a share of the haul, and so on), they may be criminally responsible for providing the code even if, unbeknownst to them, the criminal (or someone else) would have threatened to break their knees had they not provided the code. It is difficult to accept that the clerk should be exempted from criminal responsibility just because of the existence of a coercive backup mechanism of which they were not even aware. So even if criminal responsibility requires lack of acceptable alternatives rather than lack of actual alternatives, such an example, which is a variation on Frankfurt's Jones₄,⁶¹ is as challenging to the legal requirement of *acceptable* alternatives as the original Jones₄ is to the philosophical requirement of *actual* alternatives. Furthermore, if Frankfurt's examples successfully show that responsibility is possible even in the absence of actual alternatives, it is all the more so in the absence of acceptable alternatives. So, if Frankfurt's examples have been forceful enough to sway philosophers to accept that responsibility does not require actual possibilities, they should also incline legal scholars to accept that responsibility does not require acceptable alternatives. Hence, the vulnerability to Frankfurt-type examples of actual alternatives goes hand-in-hand with the vulnerability to Frankfurt-type examples of acceptable alternatives.

Conclusion

Since Frankfurt's attack on the role of alternative possibilities, the legal and philosophical discussions have progressed along strikingly different paths. While many philosophers have accepted that a person may be responsible even if they lacked an alternative, the approach of most jurisdictions and many legal scholars to the defence of duress has continued to treat lack of alternative as responsibility-negating, even when the coercion and the person's own volition coincide, as in the case of Bednarek. This paper has sought to vindicate the legal approach to alternative possibilities by identifying a principle that differs from the one Frankfurt attacked. By focusing on the epistemic element of (justified) belief in the lack of an alternative, rather than on its ontological absence, this principle was shown to be both immune to Frankfurt's attack and more suitable for grounding the legal approach. The way in which the case of Bednarek was handled by the prosecution and the court might be worthy of criticism on various grounds,⁶² but it is little surprise that none resembles Frankfurt's attack on the importance of alternative possibilities.

There are some issues involving the lack of alternative that seem to be missing from contemporary philosophical discussion of moral responsibility and alternative possibilities. Consider a case in which Jones had substantial evidence for thinking that he had no alternative but to perform the action, but the evidence

61. For another attempt to convert Jones₄ to the legal context, see Tadros' example, discussed in note 45 above.

62. See *supra* note 4.

was not conclusive. Is he responsible for performing the action? The epistemic dimension is particularly important, given the abundance of real-life cases in which the agent has some evidence about their alternatives, yet it is unclear whether the available evidence warrants justified belief.

While some philosophers might accept that the epistemic dimension is important, current philosophical debate seems to assume that a proper discussion of the epistemic dimension should be conducted only after settling the role of alternative possibilities and the conditions of responsibility. That is to say, there is an assumption that it is first necessary to determine whether the availability of an alternative is a precondition of responsibility, and only once this question is settled should the further question of how the agent's beliefs affect their responsibility be approached.

Unlike analytical philosophy, which excels in breaking complex and murky issues into well-defined questions, and carefully addressing them with hypothetical 'clean' examples, legal analysis is often burdened with the necessity of reaching concrete conclusions about real cases, which are hardly ever 'clean'. Perhaps it is this very murkiness that renders legal analysis of coercion and volition more perceptive. The need to bring together all the relevant dimensions of the case before reaching a final conclusion about the agent's (criminal) responsibility does not allow one to focus (only) on the alternative possibilities that the agent has, without taking into account their beliefs about them.

This paper has thus suggested that the philosophical discussion should be turned on its head. The question of alternative possibilities is crucial to the agent's responsibility, but mainly via their own perspective. Discussion of the agent's responsibility in cases involving a lack of alternatives should not be conducted through hypothetical and unrealistic examples such as Jones₄, but through the rich body of real cases that raise important epistemic questions about the agent's beliefs concerning the coercion to which they are exposed. Shifting the focus from the ontological to the epistemic serves as a reminder that the agent's responsibility cannot be evaluated without taking into account their epistemic position *vis-à-vis* their alternative possibilities. This shift of focus both reflects and highlights the centrality of the agent's internal point of view to the determination of their responsibility.