

Michael Otsuka

Incompatibilism and the avoidability of blame

**Article (Published version)
(Refereed)**

Original citation:

Otsuka, Michael (1998) *Incompatibilism and the avoidability of blame*. [Ethics](#), 108 (4). pp. 685-701. ISSN 0014-1704

DOI: [10.1086/233847](https://doi.org/10.1086/233847)

© 1998 [University of Chicago](#)

This version available at: <http://eprints.lse.ac.uk/54002/>

Available in LSE Research Online: November 2013

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.



Incompatibilism and the Avoidability of Blame

Author(s): by Michael Otsuka

Source: *Ethics*, Vol. 108, No. 4 (July 1998), pp. 685-701

Published by: [The University of Chicago Press](#)

Stable URL: <http://www.jstor.org/stable/10.1086/233847>

Accessed: 14/11/2013 06:29

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *Ethics*.

<http://www.jstor.org>

Incompatibilism and the Avoidability of Blame*

Michael Otsuka

In this article I address a topic that is foundational to moral philosophy: that of the conditions that must obtain if human beings are to be worthy of blame for wrongdoing. My ambition is to provide one significant part of the explanation of why no one would be worthy of blame if the universe were causally determined.¹ The most familiar argument for the incompatibility of determinism and blameworthiness can be presented in roughly the following form:

Familiar argument for incompatibilism:

1. One is blameworthy for performing an act of a given type only if one could have refrained from performing an act of that type. (I shall call this claim the 'Principle of Alternate Possibilities'.)²
2. If determinism is true, then one never could have refrained from performing acts of whatever types that one has performed.

* I thank the following people who read and provided commentary on earlier drafts: R. Albritton, T. Burge, F. Bruno, G. A. Cohen, M. Della Rocca, C. Dingman, J. M. Fischer, S. Foran, T. Hall, S. Hansen-Castro, B. Herman, A. Hsü, S. L. Hurley, M. Lange, M. McKenna, A. Rajcz, S. Shiffrin, J. Tannenbaum, M. Thau, K. Vihvelin, G. Watson, R. Wedgwood, an anonymous referee for and two anonymous editors of *Ethics*, and especially David Copp. I have also profited from discussion with members of the Law and Philosophy Discussion Group and those who attended talks at Yale; the University of California, Santa Barbara; the University of California, Irvine; the University of California, Davis; and the University of Colorado at Boulder.

1. I understand causal determinism to be the claim that "the prevailing laws of nature are such that there do not exist any two possible worlds which are exactly alike up to some time, which differ thereafter, and in which those laws are never violated." Here I follow David Lewis, "Causation," *Journal of Philosophy* 70 (1973): 556–67, p. 559.

2. Here I follow Harry Frankfurt's nomenclature. This claim is a narrower version of, and implied by, Frankfurt's unrevised version of the Principle of Alternate Possibilities, according to which one is *morally responsible* for performing an act of a given type only if one could have refrained from performing an act of that type. See Frankfurt, "Alternate Possibilities and Moral Responsibility," *Journal of Philosophy* 66 (1969): 829–39. In Frankfurt's version, the requirement also applies to acts for which one is worthy of praise and those acts for which one is held morally accountable even though neither praise nor blame is appropriate (e.g., the signing of a contract that accompanies purchases made by credit card). In this article I limit myself to a discussion of acts for which one is blameworthy.

Ethics 108 (July 1998): 685–701

© 1998 by The University of Chicago. All rights reserved. 0014-1704/98/0804-0002\$02.00

3. Therefore, if determinism is true, then one is never blameworthy for performing an act of a given type.

The second premise of this argument is controversial. But I will not discuss it here. Rather, my focus will be on the first premise. The Principle of Alternate Possibilities that constitutes this premise went largely unchallenged before the publication nearly thirty years ago of Harry Frankfurt's "Alternate Possibilities and Moral Responsibility."³ In that article Frankfurt presented an ingenious counterexample to this principle which, in the opinion of many, presents the most serious challenge to incompatibilism to date. Incompatibilists have responded to Frankfurt's challenge in a variety of ways. Some have tried to show that his example is not a genuine counterexample to the Principle of Alternate Possibilities. I do not pursue this strategy here, since I am inclined to believe that Frankfurt's example *is* a genuine counterexample to this principle. Rather, my strategy is to propose that the Principle of Alternate Possibilities be rejected in favor of a different incompatibilist principle that I call the 'Principle of Avoidable Blame'. This principle can be deployed in an argument for incompatibilism that is closely related to the familiar argument. In Section I, I demonstrate that the Principle of Avoidable Blame is resistant to counterexample of the sort that Frankfurt has shown to embarrass the Principle of Alternate Possibilities. In Section II, I present a positive argument for the Principle of Avoidable Blame that appeals to the relation of blame to the "reactive attitudes" of resentment and indignation. In Section III, I argue against the possibility of blamelessly stumbling into a "moral blind alley" where, contrary to the Principle of Avoidable Blame, one would be blameworthy for whatever one is capable of doing.

I assume throughout that, when we say that X is blameworthy for performing an act of a given type (e.g., for killing Y), we are saying that this person is blameworthy under a given description of what she has done, where this description specifies one of the types of thing that this act was.⁴ Often the description under which a person is blameworthy refers to consequences that extend beyond the movements of her body. If, for example, X moved her trigger finger, we might justifiably blame her, not just for doing that, but also for pulling the trigger and for killing Y, where the latter types of act are specified by descriptions that refer to consequences that extend beyond the movements of her body. We might, of course, be justified in blaming X for killing Y even if such a

3. *Ibid.*

4. Here I follow Elizabeth Anscombe, who introduced the locution of an act's being intentional under one description but unintentional under another. She has also written of being obliged to do something under one description but not obliged under another. See Anscombe, *Intention* (Oxford: Blackwell, 1957), p. 11, and "'Under a Description,'" *Notis* 13 (1979): 219–33.

killing was unintended—even if, for example, her only intention was to maim. But in such circumstances X need not be blameworthy under every description of what she has done that refers to an unintended consequence of the moving of her trigger finger. X's alerting her neighbor to her crime by the sound of gunshot involves an unintended consequence. But she is not blameworthy for alerting her neighbor.⁵

I

Frankfurt's counterexample to the Principle of Alternate Possibilities runs as follows.⁶ Suppose an indeterministic world in which people can normally do otherwise. Imagine that somebody in this world named Jones killed an innocent person named Smith, and that he killed him wholeheartedly, with premeditation, for selfish gain, and without any prompting.⁷ According to Frankfurt, Jones might be blameworthy for killing Smith even if he could not have refrained from doing so. For we can imagine that Jones could not have refrained for the following reason: had it become clear to somebody named Black (who is an excellent judge of such things) that Jones was about to decide not to kill Smith, then Black would have intervened and forced him to do so.⁸ But Black never had "to show his hand because Jones, for reasons of his own,

5. I believe that when X moves her trigger finger, she performs a single act of indefinitely many types (e.g., pulling the trigger, firing a bullet, killing Smith, killing a human being, killing before breakfast, alerting her neighbor, etc.) rather than many distinct acts of these different types. But nothing of substance in this article hangs on the correctness of the former rather than the latter view regarding the individuation of acts. I could reformulate, without diminishing the strength of, my argument in terms of this latter view.

6. Frankfurt credits Nozick for having made a similar point in earlier unpublished lectures. See Frankfurt, "Alternate Possibilities," p. 835, n. 2. John Locke presented a similar counterexample to the claim that voluntariness requires ability to do otherwise. See Locke, *Essay concerning Human Understanding*, bk. 2, chap. 21, sec. 10.

7. I have taken the liberty of adding a bit of detail to the example. Frankfurt does not himself specify the nature of Jones's deed.

8. Frankfurt proposes that we let "Black pronounce a terrible threat" and thereby coerce Jones into killing Smith. To those who maintain that it is impossible by such means literally to render it impossible that someone refrain from performing an action, Frankfurt proposes that we let "Black give Jones a potion, or put him under hypnosis, and in some such way as these generate in Jones an irresistible inner compulsion to perform the act Black wants performed and to avoid others. Or let Black manipulate the minute processes of Jones's brain and nervous system . . . so that causal forces running in and out of his synapses and along the poor man's nerves determine that he chooses to act and that he does act in the one way and not in any other" (Frankfurt, "Alternate Possibilities," pp. 835–36). To those who maintain that such neural manipulation is not compatible with agency on the part of Jones, we can imagine, on Frankfurt's behalf, that Black is an omnipotent being who has the power to impose deterministic laws of physics that make it inevitable that Jones kill Smith. Frankfurt's opponent would not want to deny the compatibility of determinism and action, for such a denial would beg the question against Frankfurt, since then, a fortiori, determinism would have to be false for there to be action for which one could be blameworthy.

decide[d] to perform, and [did] perform, the very action that Black want[ed] him to perform.”⁹ Moreover, Jones had no idea that Black would have intervened and forced him to kill Smith if he had not done so on his own. Given the nature of Jones’s deed, it is hard to deny that he would have been blameworthy for killing Smith if, *ceteris paribus*, Black had been altogether absent from the scene and Jones had killed Smith even though he could easily have refrained from doing so and knew that he could have refrained. Moreover, Frankfurt contends that Black’s presence should make no difference to whether or not Jones is worthy of blame. Even though Black’s presence and his readiness to intervene were sufficient to ensure that Jones could not have refrained from performing an act of the type “killing Smith,” this fact does not supply Jones with a legitimate excuse for what he has done. Frankfurt believes that this is so because Black in fact exerted no influence whatsoever on Jones’s behavior even though he rendered it impossible for Jones to have refrained from killing Smith.

In the light of this counterexample, I propose that the Principle of Alternate Possibilities be replaced by a different incompatibilist principle—the aforementioned Principle of Avoidable Blame. Unlike the Principle of Alternate Possibilities, the Principle of Avoidable Blame is sensitive to the ethical quality of—that is, one’s blameworthiness or blamelessness with respect to—one’s alternatives. According to this principle:

*Principle of Avoidable Blame:*¹⁰ One is blameworthy for performing an act of a given type only if one could instead have behaved in a manner for which one would have been entirely blameless.

To clarify this principle: (1) It merely states a necessary condition, and not a partial definition, of blameworthiness. (2) When I say that one could instead have behaved in a manner for which one would have been entirely blameless, I mean that it was within one’s voluntary control whether or not one ended up behaving that way.¹¹ But I need not claim that the behavior itself must have been voluntary. (3) I argue below that

9. *Ibid.*, p. 836.

10. A note on nomenclature: like the Principle of Alternate Possibilities, the Principle of Avoidable Blame states that the presence of an alternate possibility is a necessary condition of blameworthiness. Were it not for the fact that the Principle of Alternate Possibilities is already so well known by that name, I would have given it a name that differentiates it from the Principle of Avoidable Blame and subsumed both it and the Principle of Avoidable Blame under a genus by the name of ‘Principles of Alternate Possibilities’.

11. Suppose that someone would have behaved in a manner for which she would have been entirely blameless if and only if she had had a totally unexpected, involuntary, and incapacitating seizure. Suppose that she could have had such a seizure insofar as this was a physiologically live possibility. There is perhaps a sense in which she could have behaved in a manner for which she would have been entirely blameless. But for the purpose of interpreting the Principle of Avoidable Blame, she could not have so behaved.

one would instead have been entirely blameless if one had behaved least badly in comparison with all of one's other options.¹² (4) By "entirely blameless," I mean "blameless under any description of what one has done."¹³

The Principle of Avoidable Blame can be deployed in the following revised version of the familiar argument for incompatibilism:

Revised argument for incompatibilism:

- 1'. One is blameworthy for performing an act of a given type only if one could instead have behaved in a manner for which one would have been entirely blameless (Principle of Avoidable Blame).
- 2'. If determinism is true, then one never could instead have behaved in a manner for which one would have been entirely blameless.
3. Therefore if determinism is true, then one is never blameworthy for performing an act of a given type.

The revised second premise follows from the conjunction of the second premise of the familiar argument¹⁴ and the following claim: if one is blameworthy for performing an act of a given type, then one could instead have behaved in a manner for which one would have been entirely blameless only if one could have refrained from performing at least one type of act that one has performed. This claim is true for the following reason. Suppose that one is blameworthy for actually performing an act of a given type but would have been entirely blameless in a different scenario. It follows that one would in some respect have behaved less badly in this different scenario. This difference in behavior

12. As I shall define the notion of "behaving less badly," one would behave less badly by X-ing rather than Y-ing if and only if, given one's factual knowledge of one's circumstances, one would have stronger moral reason to X rather than Y if one could do either. In order to determine the strength of one's moral reasons, one must consider such familiar morally relevant factors as the expected harmfulness of one's behavior, the extent to which such behavior would come into conflict with one's duties and obligations, the nature of one's intentions, one's motives, and so forth.

13. One further qualification: even if one found oneself in a predicament in which one is worthy of blame for what one has done and would have been worthy of blame for whatever else one could have done in this predicament, one nevertheless could have behaved in a manner for which one would have been entirely blameless if one landed in this predicament as the result of a previous choice for which one is worthy of blame. For example: one found oneself in a predicament in which one could not have prevented one's car from barreling through a crosswalk filled with schoolchildren. But one landed in this predicament because of a previous choice for which one is blameworthy to drive far in excess of the posted speed limit. The Principle of Avoidable Blame also applies to this previous choice: one is blameworthy for it only if one could instead have behaved in a manner for which one would have been entirely blameless.

14. Recall that this premise states that if determinism is true, then one never could have refrained from performing acts of whatever types that one has performed.

can be captured in terms of one's having refrained in this different scenario from performing at least one type of act (specified at some level of description) that one actually performed.¹⁵

Given the above claim, it is a consequence of the Principle of Avoidable Blame that if one is blameworthy for performing an act of a given type, then one must have been able to refrain from performing at least one type of act that one has performed. But unlike the Principle of Alternate Possibilities, the Principle of Avoidable Blame does not impose, as a requirement of blameworthiness for performing an act of a given type, that one have been capable of refraining from performing an act of the given type for which one is worthy of blame. So long as one could instead have been entirely blameless while performing an act of this type, one can be blameworthy for performing, even if one could not have refrained from performing, an act of this type.

Some have defended the Principle of Alternate Possibilities against Frankfurt's counterexample by arguing that Jones is not blameworthy for performing an act of the type "killing Smith *simpliciter*"; rather, he is blameworthy for performing an act of the type "killing Smith *on his own*." And even though Jones could not have refrained from killing Smith, he could have refrained from killing Smith on his own. He could instead have killed Smith as a result of compulsion.¹⁶ This line of defense is controversial, since it is arguable that one needs to draw too fine a distinction in order to maintain that Jones is blameworthy for killing Smith on his own while at the same time denying that he is blameworthy for killing Smith. It is a virtue of the Principle of Avoidable Blame over the Principle of Alternate Possibilities that, even if Jones is blameworthy for killing Smith (and not merely for killing Smith on his own), Frankfurt's example does not refute the Principle of Avoidable Blame. It follows from what I say below that if Jones is indeed blameworthy for killing Smith, then he could have behaved in a manner for which he would have been entirely blameless. Jones could have behaved in such a manner even if Black's presence and readiness to intervene were enough to ensure that Jones could not have refrained from performing an act of the type "killing Smith." Hence delicate questions regarding the precise delineation of the type or types of act that Jones is really worthy of blame for performing—questions whose answers are crucial to a determination of whether Frankfurt's example refutes the Principle of Alternate Possibilities—are irrelevant to an assessment of whether Frankfurt's example refutes the Principle of Avoidable Blame.

I now explain why the Principle of Avoidable Blame is resistant to

15. I am indebted to David Copp for the ideas in this paragraph.

16. Margery Bedford Naylor offers this criticism of Frankfurt's counterexample in "Frankfurt on the Principle of Alternate Possibilities," *Philosophical Studies* 46 (1984): 249–58.

counterexample of the sort that Frankfurt has deployed against the Principle of Alternate Possibilities.

Frankfurt's counterexample to the Principle of Alternate Possibilities refutes the Principle of Avoidable Blame just in case it is an example in which Jones is blameworthy (under at least one description of what he has actually done) but could not instead have behaved in a manner for which he would have been entirely blameless. In Frankfurt's example Jones would have ended up behaving in a manner for which he would have been entirely blameless if, instead of doing what he did, he had entertained those thoughts that would have led Black to conclude that he was about to decide not to kill Smith. At this point Black would have stepped in and forced Jones to kill Smith. Recall that Frankfurt has suggested that he would have done so by pronouncing a terrible coercive threat, inducing an irresistible impulse by means of hypnosis or potion, or directly manipulating Jones's brain and nervous system. Compatibilists and incompatibilists alike would agree that each of these methods would be sufficient to absolve Jones of blame for killing Smith (and for whatever he would have done under any other level of description).¹⁷ Jones therefore had an entirely blameless alternative.

But this still leaves open the question of whether Jones *could have* behaved in a manner for which he would have been entirely blameless—that is, whether it was within his *voluntary control* that he ended up behaving this way. Frankfurt has proposed that Black's intervention would have been triggered by an involuntary twitch that Jones would have registered if and only if he was about to decide to refrain from killing Smith.¹⁸ We are to suppose that this twitch would have been caused by the sort of thought processes that would always and only have preceded a decision on the part of Jones to refrain from killing Smith.

Now the twitch would have been the result of thought processes over which Jones either had voluntary control or not.

Suppose, on the one hand, that these thought processes are something over which Jones had voluntary control. In this case Frankfurt's example would involve Jones's voluntary control over that which would have led to his doing something for which he would have been entirely blameless. The Principle of Avoidable Blame is therefore unrefuted, since Jones could have behaved in a manner for which he would have been entirely blameless.

Suppose, on the other hand, that Jones lacked voluntary control over the twitch-inducing thought processes that would have preceded any decision to refrain from killing Smith. In this case Jones would not have had voluntary control over whether, instead of doing what he did,

17. More precisely, they would agree if this claim is qualified in the manner indicated in n. 24 below.

18. Frankfurt, "Alternate Possibilities," p. 835, n. 3.

he ended up behaving in a manner for which he would have been entirely blameless because of Black's intervention. Hence he could not have behaved in a manner for which he would have been entirely blameless because of Black's intervention.¹⁹ Nevertheless Jones could have behaved less badly without provoking Black's intervention: he could have killed Smith from a nobler motive, or without premeditation, or less wholeheartedly. It follows from what I say in the next section that since Jones could have behaved less badly, he could have behaved in a manner for which he would have been entirely blameless. Hence, the Principle of Avoidable Blame is, once again, unrefuted.²⁰

One might try to modify Frankfurt's counterexample to the Principle of Alternate Possibilities so that it is more closely tailored to refute the Principle of Avoidable Blame. I do not think that any such modification would give rise to an example that refutes this latter principle. Suppose once again, for the sake of trying to construct such an example, that Jones killed an innocent person named Smith, and that he killed him wholeheartedly, with premeditation, for selfish gain, and without any prompting. If Black had been entirely absent from the scene, then it would have been within Jones's voluntary control to behave less badly. But, as before, Black is lurking in the background and monitoring Jones's behavior. In order for this example to refute the Principle of Avoidable Blame, the following must be true: had it become clear to Black (who remains an excellent judge of such things) that Jones was about to decide to behave *any less badly* than the manner in which he actually ended up behaving, then Black would have intervened to ensure that Jones ended up behaving no less badly.

It is not clear how such intervention would have succeeded.

If, on the one hand, it would have involved the bringing to bear of pressure on Jones that is supposed to make it inevitable that he end up behaving no less badly, then it would have involved a scenario in which Jones would in fact have behaved less badly, indeed would have been excused from blame, on account of his having been irresistibly pressured into doing something. Each of the aforementioned methods of intervention that Frankfurt has suggested—coercive threat, potion or hypnosis, or direct neural manipulation—would have been sufficient to excuse Jones from blame for what he ended up doing.²¹ Less intrusive means of

19. See n. 11 above and accompanying text.

20. If, however, Jones could neither have behaved less badly in any fashion that would not have provoked Black's intervention nor have entertained those thoughts that would have triggered Black's intervention, then Frankfurt is not entitled to the claim that Jones is blameworthy for killing Smith. Such a claim would beg the question against the incompatibilist, since in this case Jones could not have done otherwise in any possibly morally relevant respect even if Black had been entirely absent from the scene.

21. It would make no difference if Black were replaced by an imperceptible, impenetrable "force field" that happens to contour itself perfectly to all of Jones's actual thoughts,

bringing irresistible pressure to bear on Jones would also have been sufficient to excuse him. Suppose, for example, that, had it become clear to Black that Jones was about to decide to behave any less badly, then Black would have stepped in and tempted Jones to do something (that is at least *prima facie*) wrong and that Jones would not have been able to do other than succumb to this temptation.²² This may appear to be a case in which Black is able to ensure that Jones behave in a blameworthy fashion. But this appearance is deceptive. When one says that one could not help but succumb to temptation, one typically says something that is not strictly speaking true: such temptation, however great, is rarely literally irresistible. When the temptation is not literally irresistible, we often hold the person blameworthy. But here the person could instead have behaved in a manner for which she would have been entirely blameless, and hence the Principle of Avoidable Blame is not called into question. Only in highly extreme and unusual cases is the temptation to do wrong literally irresistible. Such cases might, for example, involve the prospect of relief from excruciating torment.²³ But in these cases the person is excused on account of the severity of the pressure that was brought to bear.²⁴ We should therefore be careful not to export our intuitions regarding blame in more ordinary cases to these extraordinary cases.

If, on the other hand, Black's intervention would not have involved the exerting of literally irresistible pressure on Jones, then such intervention could not have ensured that Jones would have ended up behaving no less badly. Hence we do not have a counterexample to the Principle of Avoidable Blame in the absence of intervention that involves irresistible pressure.

choices, and actions without influencing them at all. It might appear that this force field rendered it impossible for Jones to have refrained from doing anything that he did and hence that it rendered it impossible for Jones to have behaved any less badly than he did. This appearance is deceptive since Jones could, instead of acting independently of the force field, have voluntarily (but unintentionally) run up against this field and consequently have been irresistibly forced to kill Smith. Hence, even though he is blameworthy (under at least one description of what he has actually done), he could have behaved in a manner for which he would, as the result of such force, have been entirely blameless. (I thank an anonymous referee for drawing my attention to this case.)

22. I am indebted to John Campbell for this example.

23. We can imagine that someone presents a heroin addict in the throes of withdrawal with the easy opportunity to steal some heroin from its rightful owner.

24. In cases in which one succumbs to temptation or other pressures that irresistibly move one to act, Frankfurt maintains that one is not blameworthy if one unwillingly succumbs but that one may be blameworthy if one willingly succumbs to this irresistible pressure. See Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68 (1971): 5–20, esp. sec. 4. I maintain that one is not blameworthy for anything at all in the latter case if one had no blameless alternative to willingly succumbing to this pressure (and was not at fault for having no such alternative). I stipulate, in the cases under discussion, that if one willingly succumbed to the irresistible pressure, then both the willingness and the succumbing were made irresistible by this pressure.

Even if one manages to construct an example that overcomes these difficulties, I do not think that such an example would refute the Principle of Avoidable Blame. Let us assume, for the sake of argument, that it is somehow possible to construct an example in which, unbeknownst to Jones and without actually exerting any influence on him, Black (or someone or something else) closed all possibility that Jones have behaved any less badly than he actually behaved. It follows from what I say in the next section that, in this case, Jones would not be worthy of blame for what he has done. He would not be worthy of blame even if he would have been blameworthy if Black (or this other person or thing) had not closed all possibility that Jones have behaved less badly. It therefore follows from what I say in the next section that, contra Frankfurt, the closing of alternate possibilities can make a difference to whether or not someone is worthy of blame even if that person's behavior is entirely unaffected by the closing of these alternatives.

In the next section, I explain why I affirm the Principle of Avoidable Blame. Of particular relevance to the arguments I have advanced in this section, I argue that blaming someone for what she has done is warranted only if she could have behaved less badly and that if she could have behaved less badly, then she could have behaved in a manner for which she would have been entirely blameless. In order to do so, I first distinguish blame from something else that can genuinely be unavoidable and that others have mistaken for blame.

II

Robert Adams has defended the thesis that one might legitimately be unavoidably blameworthy for one's attitudes, temperament, or character.²⁵ If such unavoidable blameworthiness for *the way one is* were warranted, then serious doubt would be cast on my thesis that one cannot be unavoidably to blame for *what one has done*. For we would have discovered that it is not a general fact about blame that it can never be both unavoidable and justifiable. I would, however, like to affirm this general fact about blame. I grant that one's arrogance, callousness, ingratitude, tendency toward *Schadenfreude*, and so on, may have been involuntarily formed and may remain beyond one's voluntary control. I also grant that these traits of character are nevertheless properly regarded as vices. According to Adams, the attribution of a vice to somebody can properly be construed as a form of justifiable blame even if the vice is involuntary. Moreover, the blame in question is, as he would describe it, "moral" blame insofar as it is condemnation for a moral failing. I agree with Adams that the attribution of a vice to somebody is the attribution of a bad-making property that is ethical in nature, unlike the nonethical bad-making properties of stupidity, athletic ineptitude, or ugliness. But the

25. Robert Adams, "Involuntary Sins," *Philosophical Review* 94 (1985): 3–31, pp. 21–24.

attribution of a vice does not necessarily involve blame. For one can properly regard somebody as nasty and cruel while still leaving open the question of whether that person is *blameworthy* for being the way she is. One does not withdraw these attributions of vices on concluding that the person is not to blame for being this way because she is a psychopath who was deprived of oxygen in the womb and severely abused as a child. Even in the light of these findings, she is still nasty and cruel, and these are still vices of character.

When one asks whether someone is blameworthy either for a vice of character or for what she has done, one wants to know whether something more than an attitude of horror, loathing, disgust, or pity, or a policy of avoidance, management, quarantine, or elimination, is called for. One wants to know whether, in addition or instead, a “reactive attitude” of a different sort is warranted. This attitude is aptly described as indignation.²⁶ I believe, and shall assume, that someone is blameworthy for the way she is or for what she has done if and only if indignation on account of the way she is or what she has done would be warranted.²⁷ In the remainder of this section, I argue that indignation, and therefore blame, that is directed at someone for what she has done is warranted only if she could have behaved less badly.²⁸ The Principle of Avoidable Blame follows from this claim if we plausibly assume that (at least) one of the ways in which she could have behaved less badly is the least bad way that she could have behaved. She would have been blameless for what she has done if she had behaved in this way, since she could not have behaved less badly than it.

According to Peter Strawson: “If someone treads on my hand accidentally, while trying to help me, the pain may be no less acute than if he treads on it in contemptuous disregard of my existence or with a malevolent wish to injure me. But I shall generally feel in the second case a

26. Here I follow Peter Strawson, who regards indignation and resentment as impersonal and personal versions of the same attitude. On his account, indignation in response to the actions of others is a reaction “to the quality of others’ wills, not towards ourselves, but towards others.” Resentment, by contrast, is a reaction to the qualities of others’ wills toward ourselves. Indignation is therefore the “vicarious analogue of resentment”; it is, in other words, “resentment on behalf of another, where one’s own interest and dignity are not involved.” See Strawson, “Freedom and Resentment,” *Proceedings of the British Academy* 48 (1962): 187–211, pp. 199–200. I shall employ the term ‘indignation’ to encompass resentment as well as its impersonal analogue. (Strawson himself notes that his own restriction of ‘indignation’ to the impersonal is artificial, since “one can feel indignation on one’s own account,” and this is just another name for resentment [ibid., p. 200].)

27. Compare Allan Gibbard: “An observer thinks an act *blameworthy* . . . if and only if he thinks it rational for the agent to feel guilty over the act, and for others to resent the agent for it” (Gibbard, *Wise Choices, Apt Feelings* [Cambridge, Mass.: Harvard University Press, 1990], p. 47).

28. I also affirm the analogous claim that indignation, and therefore blame, that is directed at someone for the way she is (i.e., for her character) would be warranted only if she could have had a better character.

kind and degree of resentment that I shall not feel in the first.”²⁹ I would qualify this observation by adding that indignation (of this personal sort) would be warranted in the second case only if the person did what he did even though he could have behaved less badly. Moreover, indignation would be warranted in the first case if the person, while wishing me no ill, had exercised far less caution than he could and should have. Any indignation should dissipate, not only if I discover that he trod on my hand accidentally while exercising due caution, but also if I discover that he did so malevolently, or in contemptuous disregard, but could not have behaved any less badly. A person would be worthy of indignation for malevolently inflicting pain only if such infliction was gratuitous—not in the sense that it was done for no reason (he could well have had ample selfish or malevolent reason), but—in the or sense that it was an expression of the agency of someone who was free, and knew (or ought to have known) that he was free, to behave less badly instead. It is the fact that such a person behaved so badly even though he knew (or ought to have known) that he didn’t have to that makes his behavior galling and hence worthy of indignation.³⁰

Take any imagined pair of individuals who have behaved badly (e.g., who have maliciously injured another) and hold everything constant except for the fact that the one could have behaved less badly, and knew that she could have, whereas the other could not have behaved less badly. The fact that the one person behaved as badly as she did even though she knew that she didn’t have to provides sufficient grounds for indignation in her case that are lacking in the second case. Moreover, there are no other grounds that are sufficient for indignation in this second case. Such grounds are lacking no matter how malevolent or otherwise vicious this person might have been. But, one might ask, what if the second individual possessed the justified (but false) belief that she could have behaved less badly but nevertheless chose to behave badly in spite of this belief? Suppose that she could not have behaved less badly because everything about her was causally determined (assuming that determinism renders it impossible to have behaved less badly), but she didn’t know this fact and didn’t think for a minute that she had no option but to injure this person maliciously. Is it so clear that she is not to blame for what she has done just because (astonished as she would be to hear it) she could not actually have behaved any less badly?³¹

29. Strawson, p. 191.

30. But if the cost of behaving less badly was so high that it would have been superegregatory to do so, then indignation is not warranted. I shall set this possibility aside in the following discussion.

31. Here I paraphrase an objection that Rogers Albritton has offered in correspondence with me. Albritton believes that the Principle of Avoidable Blame is more plausible in cases (unlike the above) in which someone saw no alternative to doing what she did that

Although I acknowledge the skeptical force of this question, I do not think she is to blame. One's knowledge that someone justifiably (albeit falsely) believes that she could instead have behaved less badly is not enough to justify indignation. The offense must genuinely be gratuitous in the sense offered in the previous paragraph, and not merely believed by the offender to be gratuitous. "How dare you treat me this badly when you didn't have to, and you knew you didn't have to." This objection carries force and provides grounds for indignation. But a victim is not entitled to such an objection when she knows that the aggressor was causally determined or otherwise incapable of behaving less badly. Rather, the most she is entitled to say is: "I realize that you could not have behaved any less badly. But how dare you treat me this badly when you had the justified but false belief that you didn't have to." This accusation lacks force.

We would, of course, have every reason to attribute a vicious character to this person who could neither have behaved less badly nor have been any better because everything about her was causally determined. She is a worse person for injuring another in spite of her belief that she didn't have to than she would have been if she injured others only when she believed that she could not help but do so. In holding her blameless both for her actions and for her character, I do not obliterate an important ethically relevant difference between the following two sorts of causally determined individuals: (1) someone who knew the difference between right and wrong and who possessed a general ability to control her actions but who wholeheartedly injured somebody else for the sadistic thrill of it in circumstances (which were unavoidable) in which she could not have behaved any less badly, and (2) someone who inflicted an equally severe injury on another as the result of something outside the boundaries of her rational agency (e.g., a seizure, an obsessive/compulsive disorder, or a slip and a fall). One might be tempted to say that this difference involves a difference in the blameworthiness of the person. One might argue that in the first case we are justified in blaming the person for what she did, since it was an act which flowed from a vicious character, and she wholeheartedly identified with both the act and the character from which it arose;³² whereas, in the second case, the person is not to blame. Blame in the second case should be attributed, not to the person, but rather to her pathology or to her body qua physical object. But here one is employing a different and familiar nonmoral sense of blame—that of merely causal responsibility, which is the same sense we employ when we blame the faulty wiring for starting the fire. When

was less bad and was not at fault either for not seeing any such alternative or for there not being any (if indeed there was none).

32. Compare Frankfurt, "Freedom of the Will and the Concept of a Person."

that which irresistibly moves an agent to act is a vice of character with which (again, irresistibly) she identifies, it is easy to see how the attribution of causal responsibility might take on moral tones. But the morally relevant difference between the two cases is not a difference in moral blameworthiness. Rather the difference lies in the fact that the behavior was a manifestation of a vicious character in the one case but not the other. But, again, the presence or absence of a vicious character, even one that is causally efficacious, need not imply a difference in blameworthiness even if it implies a difference that is of an ethical nature.

III

The Principle of Avoidable Blame has implications that reach beyond the problem of free will. It rules out predicaments in which a person is unavoidably blameworthy even though the freedom of her will is not at issue. It rules out a predicament in which a fully competent, undeceived, and strong-willed adult who has the ability and opportunity autonomously to perform any of a diverse range of activities faultlessly stumbles into a "moral blind alley" where each option is so ghastly, tragic, or otherwise unacceptable that she would be worthy of blame for performing it.³³ On some interpretations of the myth, Agamemnon is alleged to be blameworthy for sacrificing his daughter even though he would have been blameworthy if instead he had exercised his only other option of abandoning his responsibilities as the commander of his fleet. Yet Agamemnon's alleged unfreedom from unavoidable blameworthiness has nothing to do with the freedom of his will. The alleged unfreedom from unavoidable blameworthiness featured in many modern versions of moral blind alleys also has nothing to do with freedom of the will. One who must, for example, lie in order to maintain the secrecy of that which a friend has told one in strict confidence, or who must abandon a dependent mother in order to join the Resistance, or who must kill an innocent in order to prevent many more innocents from being killed, is not necessarily afflicted with any impairment of powers of agency that leaves untouched those who do not find themselves in such binds.

Since the Principle of Avoidable Blame cannot coexist with moral blind alleys, I would like in this section to cast doubt on the existence of the latter.

In making the case for moral blind alleys, philosophers often point to cases in which feelings of guilt for certain things that are unavoidable and outside of one's control appear to be justified. One cannot, however, always slide from claims about the justifiability of unavoidable guilt to claims about the justifiability of unavoidable blame. The following three cases illustrate this point.

33. The phrase 'moral blind alley' is Thomas Nagel's. See Nagel, "War and Massacre," *Philosophy and Public Affairs* 1 (1972): 123–44, pp. 143–44.

1. One can feel guilt over the fact that one has survived and flourished even though others no less virtuous have suffered enormously. This guilt need not be over any failure to divest oneself of unjust riches. Rather, it might be traced to nothing other than the justified conviction that one is utterly undeserving of the good fortune that makes one's life much better than the lives of others who are utterly undeserving of their bad fortune. This guilt is not necessarily irrational; one really does not deserve one's good fortune. It would, however, be irrational to think that one must somehow be to blame for one's good fortune.³⁴

2. It is also not necessarily irrational to feel guilt over the unforeseeable harm one has caused through one's actions. Imagine that I instruct a casual friend to switch to a later flight because it will be more convenient for me to pick her up from the airport after rush-hour traffic has thinned out. She switches, and it crashes. Here I will undoubtedly feel guilt owing to the fact that a free action of mine figured in the immediate causal chain leading to my friend's death. So long as my guilt is not entirely a reflection of feelings that I am *to blame* for what I have done, it should not immediately be dismissed as irrational. Nevertheless, it is significant that, even though many believe that it would be perfectly natural for me to feel guilt, no reasonable person would think it justifiable to blame me.

3. Similar sorts of things can be said about the guilt one might feel over a choice that one has made even when, through no fault of one's own, every other available choice was as bad or worse. Many of the moral dilemmas discussed in the literature involve cases in which it is difficult for one to tell whether what one did was the best one could have done in a bad situation. It is easy to see how one might feel guilt in the face of such epistemic uncertainty, for here one can doubt that one has done the right thing. A better case for the opponent of the Principle of Avoidable Blame is one in which there is no question that one has done the right thing but nevertheless one feels guilt over what one has done. I am to imagine that I am the bystander at Judith Thomson's switch who has turned Philippa Foot's runaway trolley onto the one stranger instead of letting it run over five other strangers. If I were to come across the grief-stricken family of the one whom I killed, I am fairly certain that I would suffer feelings of guilt that would survive the thought that what I did was perfectly justifiable. Once again, such guilt, however natural and understandable, does not translate into the justified belief that I am to blame for what I have done.

An opponent of the Principle of Avoidable Blame might nevertheless insist that this principle should be rejected on the ground that its

34. Compare Herbert Morris, "Nonmoral Guilt," in *Responsibility, Character, and the Emotions*, ed. Ferdinand Schoeman (New York: Cambridge University Press, 1987) pp. 220–40, esp. pp. 236–37.

acceptance implies the repudiation of the undeniable fact of moral luck, by which one's degree of praise or blame may depend on factors beyond one's control. This objection can be met, since affirmation of the Principle of Avoidable Blame does not imply the wholesale repudiation of moral luck. The Principle of Avoidable Blame requires that whether or not one is blameworthy at all for what one has done be under one's control. But it does not require that the *degree* of one's blameworthiness be completely under one's control. It is consistent with this principle that one be more blameworthy if one's attempt at murder succeeds than if it fails, or if one's drunken driving results in the death of a pedestrian rather than not, even if luck is the only thing that makes the difference between one's killing someone or not.³⁵ In these cases, even though the degree of one's blameworthiness for one's actions may differ depending on external circumstances, whether or not one attempts murder or drinks and drives is still up to oneself. Hence, one could have behaved in a manner for which one would have been entirely blameless.

Evidence of moral luck is very strong in some cases, but not, I believe, in cases in which circumstances beyond one's control would make one unavoidably worthy of blame for what one has done. Consider the following case in which whether or not one is blameworthy *at all* partially depends on factors beyond one's control. Suppose that one would have become a Nazi collaborator rather than the innocent grocer that one is if one's parents had not emigrated from France to New York in 1938. I believe that one would have been worthy of blame for collaborating in France only if such collaboration were avoidable. Hence, this case does not cast doubt on the Principle of Avoidable Blame.

More troubling to my thesis are cases of the following sort: those in which it seems that one is morally compelled to take a risky course of action but in which one would also be worthy of blame if this gamble fails. One reaches the point, for example, at which it seems that one has no choice but to send in the commandos to try to free the hostages. Yet it also seems that one would be worthy of blame if the raid fails and all the hostages are killed (even though one would be worthy of praise if it succeeds). Thomas Nagel writes: "It is tempting in all such cases to feel that some decision must be possible, in the light of what is known at the time, which will make reproach unsuitable no matter how things turn out. But this is not true; when someone acts in such ways he takes his life, or his moral position, into his hands, because how things turn out determines what he has done."³⁶ A suppressed premise of Nagel's argument is that sometimes one has no viable option but to take such moral gambles. Contrary to Nagel, I maintain that if one believes that someone is

35. These examples and those that I discuss below were drawn from or inspired by Nagel. See Thomas Nagel, "Moral Luck," in his *Mortal Questions* (New York: Cambridge University Press, 1979), pp. 24–38.

36. *Ibid.*, pp. 29–30.

blameworthy for taking a gamble that has failed, then one is committed to the claim that this person ought to have refrained from taking this gamble. She should have refrained, not given the knowledge of hindsight, but given the facts available to her at the time of her decision. But it is surely incorrect to maintain of every gamble that fails that it should not have been wagered, given the facts available at the time of the decision. If, however, one believes that, given the facts available at the time, she *ought* to have chosen to take a certain gamble, then the fact that this gamble is obligatory immunizes the gambler from blame for any bad consequences that ensue. I believe that the same holds for morally permissible but nonobligatory gambles.³⁷

IV

I hope through my articulation and defense of the Principle of Avoidable Blame to have identified and verified the authenticity of an overlooked but reputable source of the undeniable appeal of the claim that determinism is incompatible with blameworthiness.

37. More difficult are cases in which one cannot tell, at the time of choosing, whether the gamble is impermissible, permissible, or obligatory. I believe that in these cases the uncertainty should transfer to an assessment of blameworthiness in the event that the gamble turns out a failure.