

Responsibility, Reason, and Irrelevant Alternatives

Susan Hurley

Philosophy and Public Affairs, vol. 28 (2000), pp. 205-241.

See revised version in my Justice, Luck, and Knowledge (Harvard University Press 2003)

Abstract:

This article compares two prominent reason-based accounts of responsibility, those of Wolf and Fischer. An intuition about responsibility, the *irrelevant alternative intuition*, emerges from this comparison. This intuition is a significant generalization of the intuition evoked by Frankfurt cases, and applies to a wider range of cases. According to this intuition, the outright possibility of doing otherwise, all else constant, is not what matters for responsibility. Rather, what matters is whether someone would have done otherwise, *if she could have*. If she *wouldn't* have done otherwise even *if she could have*, then it is irrelevant whether or not she could have. Fischer's mechanism-based actual sequence account of responsibility suggests one way of explaining this intuition. The irrelevant alternative intuition is illustrated by various cases involving weakness of will, evil, deprived upbringing, and medical problems.

- Contents:*
1. Background: Conditional Analyses, Frankfurt Cases, and the Distinction between Actual and Alternate Sequences
 2. Reason-based Accounts of Responsibility: Wolf and Fischer
 3. Varieties of Responsiveness to Reason and Asymmetrical Responsibility
 4. Reason-based Accounts of Responsibility: Actual Sequence vs. Alternate Sequence Principles, and the Irrelevant Alternative Intuition
 5. More on How the Irrelevant Alternative Intuition Generalizes the Point of Frankfurt Cases
 6. A Remaining Issue: Reason-responsiveness vs. Regression

1. Background: Conditional Analyses, Frankfurt Cases, and the Distinction between Actual and Alternate Sequences.

The traditional debate about free will and determinism presupposes that responsibility for what you do requires the ability to do otherwise. Compatibilists argue that the ability to do otherwise is compatible with determinism, and incompatibilists deny this.

Some compatibilists have defended their thesis by giving a *conditional analysis* of the ability to do otherwise. On this account, to say that you could have done otherwise is just to say that you would have done otherwise, if you had chosen to. For example, if the reasons for acting had been different, you might have chosen to act differently, and so have done otherwise. Your having such a conditional disposition, to act otherwise under counterfactual suppositions, is compatible with it being causally determined that you act as you do under the conditions that actually obtain.

The conditional analysis is incorrect. The ability to do otherwise entails the *outright possibility* of acting otherwise: it entails that there is a causal possibility of acting otherwise, holding all else constant. A counterfactually conditioned disposition to act otherwise is not the same thing as an outright possibility of acting otherwise, all else constant. That the former is compatible with determinism does not entail that the latter is. It might not be possible for you to choose to do otherwise, hence not be possible for you to do otherwise, holding all else constant, yet still be true that if you had chosen to do otherwise, you would have done otherwise. For present purposes we can agree that incompatibilists are correct to claim that ability to do otherwise is incompatible with determinism.

The essential point against the conditional analysis turns on the contrast between an the outright possibility of an *alternate sequence* of events, and the dispositional features of the *actual sequence* of events leading to your action (Fischer 1994, ch. 7, 8; see also Klein 1990). The ability to do otherwise requires the outright possibility of an alternate sequence of events. By contrast, your acting on the basis of a disposition that would have led you to act differently under counterfactual suppositions is a feature of the actual sequence of causes that lead to your act. The actual causes of your act, in the actual sequence of events leading to your act, may involve certain dispositional properties, so that if certain counterfactual

conditions were to obtain, then you would act differently, as the conditional analysis specifies. But your having a disposition to act otherwise under different conditions does not entail the unconditioned possibility of your acting otherwise as things are.

Conditions of responsibility can be categorized as alternate sequence conditions or as actual sequence conditions, depending on whether they require the outright possibility of an alternate sequence of events with certain features or rather require that the actual sequence of events has certain features. The alternate sequence/actual sequence distinction helps to get and keep clear distinctions between different principles of responsibility that have too often been run together. For example, the alternate sequence *could-have-done-otherwise principle* of responsibility should be distinguished from the actual sequence *regression principle* (see and cf. Klein 1990). The latter says, roughly, that to be responsible for something you must be responsible for its causes. It does not specify what substantive conception of responsibility this structural principle is applied to, but just imposes the structural condition. Applied recursively, this principle requires responsibility to reach back through the actual sequence of causes. This is incompatible with at least some substantive conceptions of responsibility. A regressive conception of responsibility may be incompatible with determinism even if the could-have-done-otherwise principle is dropped.

There are other candidate principles in each category to consider as well, which will be examined in what follows. The *could-have-avoided-blame principle* is an alternate sequence principle that may be more plausible than the could-have-done-otherwise principle (Otsuka 1998). And some version of a *reason-responsiveness* principle may provide an actual sequence dispositional account that is more attractive than regressive conceptions of responsibility (Fischer 1994, Fischer and Ravizza, 1998; cf. Wolf 1990; Wallace 1996).

I have explained how the distinction between alternate and actual sequences cuts against conditional analyses and hence against compatibilists, so far as the traditional debate goes. However, this distinction is a double-edged sword. It also makes available challenges to the presupposition of the traditional debate, namely, that responsibility does require the outright possibility of acting otherwise. Harry Frankfurt broke ranks by denying this requirement in an influential series of articles (1969, 1971, etc.). His essential

insight was in effect that responsibility turns on the character of the actual sequence of events leading to an act, not on the possibility of alternate sequences. This opens up the possibility that determinism may be compatible with responsibility, even if it is incompatible with the ability to do otherwise. Even if conditional analyses do not give correct accounts of the ability to do otherwise, some conditional analysis might give a correct account of responsibility.

In Frankfurt's well-known example (1969, 835), Black wants Jones to perform a certain action, and is prepared to go to considerable lengths to ensure this. But Black also wants to avoid showing his hand unnecessarily. Black waits until Jones is about to make up his mind and does nothing unless Black, who is an excellent judge of such things, judges that Jones is not going to do what Black wants him to do. Only if Black so judges does he intervene and take effective steps to ensure that Jones decides to do, and does, what Black wants him to do. Perhaps Black manipulates Jones' brain processes, for example. Whatever conditions are needed for it to be the case that Jones cannot do otherwise, Black makes it the case that those conditions prevail. But Black never actually has to show his hand, because Jones, for reasons of his own, decides to do and does do the very thing Black wants him to do. Frankfurt claims that in this cases the fact that Jones could not have done otherwise is irrelevant to his responsibility for what he does.

Various cases that illustrate Frankfurt's point have a certain common structure. Suppose someone acts in a prima facie blameworthy way, for bad reasons but reasons that are her reasons (Fischer and Ravizza 1991, 258). Standing by is a counterfactual intervenor who wants her to act just as she acts, but who is prepared, in case he predicts she will waver, to intervene to bring it about that she goes ahead. The counterfactual intervenor is usually personified, but it could be an impersonal force that acts as a fail-safe device. In fact no intervention is needed and none occurs; the agent does not waver but acts of her own initiative. In a case with this structure, the agent could not have done otherwise. There is an alternate possible sequence, but in it she does the same thing. Despite that, the character of the actual sequence of events leading to her act is such that she is responsible. The parallel point can be made for praiseworthiness. Cases with this structure can be called *Frankfurt cases*.

Various attempts have been made to avoid these conclusions: by distinguishing act tokens from act types, by appealing to a ‘flicker of freedom’ in the possibility of wavering, by appealing to the possibility of avoiding blame rather than the possibility of doing otherwise as a condition of responsibility, and so on. These avoidance possibilities have been widely and ably discussed in the literature, and will not be rehearsed or assessed in general terms here.

Rather, the current aim is to present a different strategy for arguing for a similar conclusion, that responsibility does not require alternate possibilities. By this means, the intuition to which Frankfurt cases appeal can be significantly generalized, and evoked from a wider range of cases, which do not involve science-fiction fail-safe devices to guarantee that the agent could not have done otherwise. The more general intuition is here called the *irrelevant alternative intuition*. This alternative strategy displays the essential commitments of the irrelevant alternative intuition, though these commitments are not fully defended here. Even if this strategy it does not ultimately make the irrelevant alternative intuition irresistible, it may be useful to display a different route to it and clarify some of the surrounding landscape of issues.

This more general intuition will be developed here by applying the distinction between actual sequence and alternate sequence principles of responsibility to two prominent reason-based accounts of responsibility, those of Susan Wolf and of John Martin Fischer. Many philosophers, from Spinoza, Kant, and Hegel¹ through to the present, have conceived of freedom in terms of rational necessity, or otherwise found close connections between rationality on the one hand and freedom and/or responsibility on the other. Wolf’s and Fischer’s reason-based views of responsibility can be seen as part of this broad tradition.² But for present purposes their reason-based accounts of responsibility are used as examples, to illustrate how issues about actual sequence vs. alternate sequence principles of responsibility generate the irrelevant alternative intuition. The latter is the real topic of my argument, rather than reason-based accounts of responsibility per se.

In this section I’ve sketched the context in which the actual sequence/alternate sequence distinction emerges. It operates as the pivot between the pre- and post-Frankfurt stages of debate about responsibility: both to undermine conditional analyses in the

traditional debate and to underwrite Frankfurt's rejection of the could-have-done-otherwise principle, which the traditional debate assumed. Against this background, let's examine how the actual sequence/alternate sequence distinction sheds light on some of the forms that reason-based accounts of responsibility can take, and in particular how the irrelevant alternative intuition emerges from the comparison of Wolf and Fischer.

2. Reason-based accounts of responsibility: Wolf and Fischer

Wolf and Fischer both base their accounts of responsibility on the idea of responding to reasons, but they do so in different ways. Wolf (1990) explains the importance for responsibility of the ability to respond to reasons. But her account does not explicitly address Frankfurt cases, nor does it clearly distinguish actual-sequence and alternate-sequence principles of responsibility in the way that Fischer (1994, etc.) does.³ To assess the impact of Frankfurt cases on reason-based views of responsibility, we can compare Fischer's actual sequence reason-based account with Wolf's reason-based account, which has alternate sequence as well as actual sequence aspects, and consider the issues between them.

Wolf sets up a three-way contrast between the Autonomy View, the Real Self View, and the Reason View. She argues for the Reason View.

The Autonomy View requires that to be responsible the agent must not only control her behavior via her will, but must have *ultimate control*: "her will must be determined by her self, and her self must not, in turn, be determined by anything external to itself" (Wolf 1990, 10). The self that governs her will must be governed by her deeper self, and so on ad infinitum. If external forces such as those of heredity and environment determine her will, then her control is only intermediate and superficial. But it is no better if "the agent's control is controlled by nothing at all", or can be explained only in terms of random occurrences or brute inexplicable facts. The idea of an autonomous agent as a prime mover unmoved, whose self can endlessly account for itself and its behavior, can seem incoherent or impossible, even if necessary for responsibility (1990, 13-14). "Ultimate control" appears to be a kind of regressive control. So the Autonomy View appears to be committed to the regression principle (though Wolf does not put it in these terms).

By contrast, the Real Self view denies that a responsible agent must be endlessly accountable to herself. In effect, it denies that responsibility must be regressive in the way the Autonomy View requires. Rather, it requires for responsibility that the agent's behavior be attributable to her real self: that she is able to govern her behavior on the basis of her will and able to govern her will on the basis of her valuational system. But it does not matter, on this view, where her real self comes from (Wolf 1990, 33-35).

We can locate the contrast between the Autonomy View and the Real Self View within the space of actual-sequence principles of responsibility. While the Autonomy View requires that control extend all the way back through the actual sequence of causes, the Real Self View is satisfied if the actual sequence leads back to the Real Self in the right way, even if the sequence does not stop there.

Wolf rejects the Real Self View. We can question the responsibility of a fully developed agent even when she acts in a way that is clearly attributable to her, because we can question whether she is responsible for her real self (1990, 37). For example, some forms of mental illness may raise this concern, or cases of deprived or traumatic childhoods. These are among the cases that lead some to think we are committed to a regressive conception of responsibility.⁴ Wolf finds an unbridgable gulf between her own view and the view that in such cases people can be wholly responsible for their actions, even though they are not responsible for the real selves from which their actions result (1990, 38-39). The claim that "To be a responsible agent is simply to be, as it were, a fully formed bad act-maker" seems to confuse the mere causal responsibility of the real self with its moral or deep responsibility (1990, 40-43; see also Otsuka 1998).

Wolf's grounds for rejecting the Real Self View may seem to favor the Autonomy View, or to commit her to a regressive conception of responsibility. However, Wolf also criticizes the Autonomy View. On the basis of her criticisms of both the Real Self View and the Autonomy View, she develops a third view, the Reason View, which she favors. The Reason View does not respect the regression principle.⁵ It is a tricky matter to locate it in relation to the distinction between actual-sequence and alternate-sequence accounts, as we shall see, since its alternate-sequence requirements are asymmetric as between praiseworthy and blameworthy acts.

The Autonomy View is committed not just to a regressive conception of responsibility, but also to the could-have-done-otherwise principle (Wolf 1990, 68). (This is not surprising; it is usual for the two principles to go hand in hand and unusual for them to be distinguished, as in Klein 1990.) Wolf's critique of the Autonomy View turns on its commitment to a could-have-done-otherwise condition for praiseworthiness. An agent who cannot help choosing what she has most reason to choose is not autonomous. Autonomy requires radical freedom: that no basis for choice necessitates, that the agent have the ability to make choices on no basis even when a basis exists (1990, 54-55). The autonomous agent who does the right thing could have done otherwise, even holding all the reasons there are for acting constant.

Wolf argues that this ability to be irrational is not desirable. Acting for a reason, in order, say, to save a drowning child, is no *more* praiseworthy if an agent could have declined to act in this way, all else constant. Acting nonautonomously but for such a reason deserves praise at least as much acting the same way but autonomously (1990, 60-61). Wolf considers the way in which skeptics about values and reasons for acting may think autonomy valuable, in the absence of any right answer to questions about what should be done, because the choices of an autonomous agent are more deeply hers, freer of bias and conditioning (1990, 62-63). But, argues Wolf, unless we are such skeptics, what we should and do value is not radical freedom but the moderate freedom to act in accordance with reason. According to the Reason View advocated by Wolf, the valuable kind of flexibility is not the ability to act despite reason, but rather is built into the ability to act in accordance with reason. Someone who acts in accord with reason can be responsible even if she could not have done otherwise (1990, 63-66).

On the Reason View, what matters for responsibility is the ability to act in accordance with reason, not the ability to act in some other way. This means that whether responsibility requires that the agent could have done otherwise depends on whether she actually acts in accordance with reason or not. Praise and blame are not symmetrical with respect to alternate possibilities. If the agent does the right thing for the right reasons, then the requirement of ability to act in accordance with reason is satisfied, and she need not have been able to act otherwise, to act irrationally, in order to be responsible for what she does.

Praise does not require alternate possibilities. As Wolf says, the ability to act in accordance with reason is compatible with the inability to do anything else (1990, 69). But if in fact the agent does not act in accordance with reason, it is then critical that she have been able to do so, to act otherwise. To be responsible for what she does when she does not act in accordance with reason, it must have been possible outright for her to act in accordance with reason. Blame does require alternate possibilities.

Compare the conditions for praise and for blame on the Autonomy View, on the Real Self View, and on the Reason View. On the Autonomy View, both praise and blame require alternate possibilities. On the Real Self View, so long as the real self is the actual source of the act, it doesn't matter whether there are alternate possibilities. Since real selves can be the sources of bad acts as well as good acts, praise and blame are also on this view symmetrical with respect to alternate possibilities. Behavior governed by a real self is behavior governed by certain dispositions, to act in certain ways under certain conditions and in other different ways if conditions were otherwise, in accordance with the agent's character, values, and desires or subjective reasons. But such conditional dispositions are not the same thing as the outright ability to do otherwise, all else constant.

On the Reason View, the regress of responsibility stops not with the agent's real self, her will and valuational system, or her merely subjective reasons for acting: the beliefs and desires she has that rationalize certain acts. Rather, it stops with what reasons there are, objectively. Someone can have beliefs and desires that rationalize a certain act, even though there are no objective reasons to do it: say, because her beliefs are false, or because her desires are out of line with objective reasons. But if objective reasons are the actual source of the act, if the agent does the right thing for the right reason, it doesn't matter if she is not responsible for those reasons or her responsiveness to them. *Why in general should we want to be responsible for the objective reasons to which we respond, anyway? Could they be objective if we were? Nor does it matter whether there are alternate possibilities. The agent who does the right thing for the right reason is praiseworthy. This is the nonregressive, actual-sequence component of Wolf's Reason View. The claim that alternate possibilities are irrelevant to the praiseworthiness of right acts done for the right reasons can be regarded as*

expressing a restricted, asymmetrical form of the irrelevant alternative intuition. Like the Real Self View and unlike the Autonomy View, the Reason View is liberal with praise.

However, the objective normative commitments of the Reason View make it asymmetrical with respect to praise and blame. If reason cannot be the actual source of wrong acts or acts not done for the right reason, then when the agent does not do the right thing for the right reasons, the conditions of responsibility on the Reason View are not necessarily met. Suppose that when we look at the actual sequence of causes of such acts, we do not find objective reasons.⁶ Then the regress of responsibility cannot stop with objective reasons. When wrong acts are in question, the ability to act in accord with reason requires the outright ability to act otherwise, all else constant: to act in accordance with reason rather than against it. Blame requires alternate possibilities (1990, 79): this is the alternate-sequence component of the Reason View. Like the Autonomy View and unlike the Real Self View, the Reason View is not liberal with blame.

The difference between the Real Self View and the Reason View turns on the difference between subjective and objective reasons. In Wolf's terms, according to the Real Self View someone is responsible if and only if she is able to form her actions on the basis of her values, whereas on the Reason View someone is responsible if and only if she is also able to form her values on the basis of the True and the Good (1990, 75). Deprivation and other unfortunate circumstances may disconnect real selves from objective reasons, and so explain, on the Reason View, why blame is inappropriate. The difference between responsibility and the lack of it is less a metaphysical difference than a normative difference, a difference in an agent's connectedness to objective reasons (1990, 70-71, 76), which may result from the difference between a good upbringing and a deprived one.

While criticizing the Real Self View, Wolf appears to be sympathetic to a regression requirement. But in developing the Reason View, she makes it clear that praiseworthiness turns on the ability to act in accord with reason, and is not undermined by failure to meet a regression requirement. So Wolf's position suggests a general line of response to the challenge to find some principle other than a regression principle that explains intuitions about blame. Perhaps blameworthiness also turns on the ability to act in accord with reason, not on a regression principle. However, Wolf does not formulate her reason-based account

explicitly in terms of the distinction between the actual sequence and alternate possibilities, or by reference to Frankfurt cases.⁷ Fischer does. His position helps to see how a reason-based view of responsibility might be able to respond to this challenge.

Fischer suggests that we view Frankfurt cases as cases in which the actual mechanism on which the agent acts is at least weakly or loosely responsive to reasons, though the manipulative mechanism that would operate in the alternate, counterfactual-intervention sequence is not responsive to reasons. By the “mechanism” on which the agent acts, Fischer means the process that leads to the action, or the way the action comes about. In the case of Jones and Black, for example, when Jones acts for his own reasons, he acts on one kind of mechanism. When he does the same thing but because his brain has been manipulated by Black, he acts on a different kind of mechanism. If Jones acted because of a hypnotic suggestion implanted by Black, he would be acting on yet another different kind of mechanism. Fischer describes some mechanisms in terms of (reason-responsive) deliberation or practical reason, and others in terms of physical processes, such as (non-reason-responsive) central nervous system processes that might realize the irresistible urge to take a drug on which an addict acts (1994, 173-174; 1998, 42, 46ff). Reason-responsive mechanisms may be reflective, as in practical reasoning, but need not be. Fischer gives no general account of when the kinds of mechanism on which agents act are the same or different, but regards it as intuitive in many cases. (See the end of section 3 for further comments about this notion of “mechanism”.)

On Fischer’s view, so long as there is no responsibility-undermining factor in the actual sequence, the agent may be responsible. It doesn’t matter if there is a responsibility-undermining factor in the alternate sequence, so long as it is not actually causally operative. Differences in responsibility depend or supervene on the actual history and causes of action, not merely on what nonactual sequences might have been like (1994, 158; see also note 2 above). What matters is the character of the actually operative mechanism of action, including its dispositional character, in particular, its responsiveness to reasons. Fischer makes a point similar to Wolf’s point against the Real Self View: it is not enough for responsibility that there is a mesh between the agent’s preferences and values, or first-order and higher-order preferences. We need to know more about how this mesh was actually

produced, since it is possible for such a mesh to be produced by a bad mechanism, which is not even loosely responsive to reasons (1987; 1994, 208).

Responsiveness to reasons and hence responsibility are compatible with determinism, even though the outright ability to do otherwise, all else constant, is not compatible with determinism. Fischer calls this combination of views “semicompatibilism” (1987, 1994). Reason-responsiveness requires that the mechanism that operates in the actual sequence have certain dispositional features: holding the mechanism constant, the agent would have responded differently in at least some circumstances in which the reasons were different. Fischer (1994) requires for blame only what he calls ‘weak’ reason-responsiveness and what is here called ‘loose’ reason-responsiveness.⁸ That is, he requires only that the mechanism on which the agent acts would, in *some* (not all) possible worlds in which there is reason to do otherwise, lead her to act on that reason and hence to do otherwise. For example, someone who shoplifts although there are sufficient prudential reasons not to, might not do so if she thought her daughter would find out. Thus Fischer does not require that the agent have an outright ability to act differently, holding all else, including whatever reasons there are, constant.⁹ Fischer’s condition of weak reason-responsiveness does not support the asymmetry between praise and blame that Wolf defends; neither praise nor blame requires outright alternate possibilities. Fischer (1994) holds that the only plausible threat to moral responsibility from determinism is the threat to alternate possibilities. If Frankfurt cases demonstrate that alternate possibilities are not required for responsibility, then determinism is no longer a threat to responsibility.¹⁰

However, to remove the threat of determinism we need to do more than register that the actual sequence is what matters. In principle, some actual sequence principles of responsibility may be compatible with determinism, but not others. Cases of hypnotism, brain tumor, and so on, lead some to endorse a regressive conception of responsibility that is arguably incompatible with determinism.¹¹ Fischer accounts for some of these cases instead in terms of the failure of reason-responsiveness.

Fischer’s 1994 version of a reason-based view of responsibility differs from Wolf’s in several ways. First, Fischer appeals to the dispositional character of the mechanism on which the agent acts in the actual sequence, namely, its reason-responsiveness, rather than

to outright alternate possibilities. Wolf appeals, by contrast, to the agent's ability to act in accord with reason. This requires there be outright alternate possibilities when the agent does not act in accord with reason, though not when the agent does do the right thing for the right reason. While in what follows we will mainly be concerned with the actual sequence/alternate sequence distinction, it is important to be clear about the way the idea of ability to act in accord with reason cuts across that distinction, while the idea of reason-responsiveness is located on actual sequence side of the distinction. Second, Fischer requires only a loose link to reason for responsibility. On his 1994 view, someone who does not act in accord with reason can be responsible in virtue of acting on a mechanism that is loosely disposed to respond to reasons: that responds to reasons in some possible world. If Fischer were instead to require a tight link, then someone who did not act in accord with reason would not be responsible. For Wolf, by contrast, when an agent does not act in accord with reason, he may still be responsible, but only if there is an outright possibility of his doing the right thing for the right reason, in this very world. Third, Wolf's account clearly invokes objective reasons, the ability to form values on the basis of the True and the Good. The ability to act in accord with reason that Wolf requires is the ability to act in accord with the reasons there are, not just the reasons the agent takes herself to have. Fischer's position on objective vs. subjective reasons is not as clear as Wolf's.¹² But the net effect of his condition of weak reason-responsiveness is that praise and blame are not asymmetric on Fischer's view, as they are on Wolf's view.¹³

The upshot of this section is this three-fold comparison of Wolf and Fischer. In the next section we will use these three issues, and the resources provided by Wolf and Fischer, to articulate and explore the space of reason-based accounts of responsibility. To summarize, the three issues that have emerged from our comparison are: (1) Is the link to reason required for responsibility tight or loose? (2) Is it a link to objective reasons, or will subjective reasons do? (3) Does the link to reason needed for responsibility impose alternate sequence demands, or can it be adequately characterized in terms of counterfactual conditionals expressing the dispositions of the mechanisms that operate in the actual sequence?

3. Varieties of responsiveness to reason and asymmetrical responsibility

First, following Fischer, we should distinguish tight and loose reason-responsiveness.¹⁴ Both are counterfactual characterizations of the mechanism that operates in the actual sequence of causes leading to someone's act. If the causes of someone's act are tightly responsive to reasons, then if there were reason to act differently and the same mechanism were operating, she would act as reason requires. A *tightly reason-responsive* mechanism of action tracks reason. Someone who does not do the right act for the right reason is not acting on a mechanism that responds tightly to reason, that tracks it closely, since there is a reason to do otherwise and she does not act on it. But even so the causes of her act might be *loosely reason-responsive*. This requires only that there is *some* possible world, which need not be close to the actual world, in which there is reason to do otherwise and the same mechanism operates and it leads her to act on that reason and hence to do otherwise. For example, what actually leads her to steal a book, wrongly and imprudently, despite sufficient reasons not to, might be responsive to at least some incentives not to steal (cf. Fischer 1994, 167; 1987). A loose reason-responsiveness requirement is, roughly, what lawyers are getting at when they ask whether a defendant would have broken the law if there had been a policeman at his elbow. But it may rule out very little other than cases in which reason gets no purchase at all on the operative mechanism, such as cases of reason-blindness, sub-rationality, or mental illness.

Weakness of will is compatible with loose reason-responsiveness but not tight reason-responsiveness. Suppose I go to the movies for a good reason. But, holding the deliberative mechanism on which I act constant, I would have gone even if I'd had a publication deadline that provided a better reason not to go. That is, I would have been weak-willed under those circumstances. The mechanism on which I act is not tightly reason-responsive. But, still holding the mechanism constant, I would not have gone to the movies if I would have lost my job by going, or if tickets had cost a thousand dollars. So the causes of my act are loosely reason-responsive. If such loose actual-sequence reason-responsiveness is sufficient for responsibility, then I would be responsible in the case of weak will. But if tight reason-responsiveness is required, someone who is weak willed is not responsible.¹⁵

Tight responsiveness to objective reasons provides a condition of responsibility that is maximal in two dimensions. Thus it can be weakened in two ways. It can be loosened, as we've seen above. Or, it can be subjectified: responsiveness could be required merely to subjective reasons, rather than to objective reasons. So there is logical space for at least three more kinds of reason-responsiveness: loose responsiveness to objective reasons, tight responsiveness to subjective reasons, and loose responsiveness to subjective reasons.¹⁶

Someone who had a bad or deprived upbringing might fail to satisfy a requirement of objective reason-responsiveness but not a requirement of subjective reason-responsiveness. For example, a deprived upbringing might result in someone with evil motivations, who acts for evil reasons, in order to hurt others. These are subjective rather than objective reasons, but they might be tracked very tightly. Agents can be more or less tightly disposed to respond to their own desires or subjective reasons as well as to objective reasons. So the mechanism on which an agent acts should not be thought of simply as his desires. Rather, the mechanism on which an agent acts may be disposed to respond to various reasons, which may include desires, more or less tightly. If subjective reason-responsiveness is sufficient for responsibility, then someone who acts on dispositions that track evil reasons tightly can be responsible. But if objective reason-responsiveness is required, and assuming that there are no objective reasons to do evil and that there are objective reasons not to do evil, then someone who acts on a mechanism that is tightly responsive to evil cannot be responsible. However, someone who is less tightly responsive to objective reasons might act for an evil subjective reason: he might act on a disposition to injure someone a little, even though he would avoid acting to injure a lot. If only loose responsiveness to objective reasons is required for responsibility, then it is possible to do a bit of evil responsibly. (It may be regarded as odd to allow that someone can do a bit of evil responsibly, but not be responsible if he is tightly responsive, or dedicated, to evil.)

Our choices between objective and subjective, and between tighter and looser, versions of reason-responsiveness will correspond to whether we hold agents with various combinations of defective upbringing and weak will responsible. If we want to allow for responsible weakness of will, we should require for responsibility something looser than tight reason-responsiveness. If we want to allow for that it is possible to be dedicated to evil

yet responsible, we should go for subjective reason-responsiveness. But if instead we want only to allow the possibility of doing an occasional bit of evil responsibly, loose objective reason-responsiveness will do the trick.

Notice that all these varieties of reason-responsiveness in themselves impose actual-sequence principles of responsibility. What is required for responsibility is that the agent actually act on a mechanism that has certain dispositions to respond to reasons under counterfactual conditions, characterized in one of these ways. The actual causes of her acts must involve certain conditional dispositions to act on reasons. This is not in general, as Frankfurt cases illustrate, the same thing as an outright possibility of acting otherwise, given the reasons there actually are. As Fischer might put the point, so long as in the actual sequence the agent acts on an appropriately reason-responsive mechanism, it does not matter what other mechanism might have operated instead.¹⁷ So, in a Frankfurt case, it does not matter if there is a counterfactual intervenor or fail-safe mechanism in an alternate sequence so that a could-have-done-otherwise principle is not satisfied. What matters is the dispositional character of the operative mechanism.

However, in the maximal case where tight objective reason-responsiveness is required for responsibility, it is hard to distinguish an actual-sequence requirement for blame from an alternate-sequence requirement. Praiseworthy acts can meet the requirement for responsibility that the actual causes of an act must respond tightly to objective reasons. However, any failure to do the right thing for the right reasons indicates that the agent is not acting on a mechanism with this actual-sequence disposition. It is not relevant here to ask how the agent would have acted under different conditions; we are concerned with how she acts under the reasons that do obtain. She cannot be acting on a mechanism of the required kind, since the objective reason is there and she does not act on it. It is not possible, on such a condition, to be blameworthy. The defect is in the actual sequence: in the disposition of the operative mechanism. But this is immediately tied to the failure to do otherwise given the reasons there were.

This maximal stance in the logical space of varieties of reason-responsiveness ties the actual sequence responsiveness requirement closely to an alternate sequence, could-have-done-otherwise requirement. This may partly explain why, although Wolf's view expresses

the irrelevant alternative intuition for praiseworthiness, nevertheless in developing her position Wolf ignores Frankfurt cases involving blame. It may also partly explain why, though she requires ability to respond to reason rather than actual responsiveness, she does not find it worthwhile to distinguish explicitly alternate-sequence from actual-sequence principles. Maximal reason-responsiveness, like Wolf's ability to act in accord with reason, generates an asymmetry between blame and praise. Blame does require an alternate sequence, even though praise does not.

Are praise and blame asymmetrical in this way? How do Frankfurt cases bear on reason-based views of responsibility and on this purported asymmetry? To address these questions we need to press the distinction between actual-sequence and alternate-sequence formulations of a reason-based view of responsibility. We compare the outright ability of an agent to act in accord with reasons with the conditional disposition to respond to reasons of the mechanism on which she actually acts. We want to consider how Wolf's ability condition of responsibility behaves in relation to a responsive-mechanism condition, and especially to examine cases in which the two conditions come apart.

Fischer and Ravizza contend that we learn from Frankfurt-type cases that a mechanism-based approach to responsibility is more useful than an agent-based approach. In such cases, the agent is not reason-responsive, because she could not have done otherwise, nor is the mechanism that would operate in the alternate sequence. But the operative mechanism is reason-responsive. By "mechanism", they explain, they do not mean anything over and above the process leading to the relevant upshot, or the way the action comes about. Actions can come about in importantly different ways. It is crucial to distinguish between the way an action comes about in the actual sequence and the way it would come about in alternate sequences. But they confess they do not have any general formula for specifying when these are the same ways or different ways. Rather, they rely on relatively clear intuitions about sameness and difference of ways actions come about, or mechanisms on which agents act. Nor do they specify in a general way how to determine which mechanism is the relevant operative mechanism; again, they simply presuppose that there is an intuitively natural choice. If cases where intuitions are about mechanism-individuation are unclear are also cases where intuitions about responsibility are unclear,

then their approach will reflect this pretheoretical fuzziness (Fischer and Ravizza 1998, 38-40, 40n, 46-47; see also Wallace 1996, 190n on mechanism worries).

In the next section, Fischer's conception of acting on a reason-responsive mechanism will be taken over and employed in an intuitive way, which is the way Fischer employs it. Issues about how mechanisms are to be individuated, or how we know what to hold constant in considering other possible worlds and the agent's responses to reasons in those worlds, are not addressed. Fischer admits he does not resolve these issues, and they cannot be resolved here. Rather, the line taken in section 4 is conditional: to the extent the mechanism approach can be defended, we can argue as in the following section for a generalized irrelevant alternative intuition. The generalized irrelevant alternative intuition as developed in section 4 below is thus hostage to making good the mechanism approach, or another approach that could do the equivalent work. One way to attack the irrelevant alternative intuition, then, would be to attack the mechanism approach and to argue, furthermore, that the work it needs to do cannot be done. The essential work that needs to be done to support the irrelevant alternative intuition, as we shall see, is to underwrite the intuitive supposition that an agent would not have done otherwise, even if she could have. In section 5 I shall go on to suggest how this supposition might be explained without relying on the mechanism approach.

4. Reason-based views of responsibility: actual sequence vs. alternate sequence principles, and the irrelevant alternative intuition

Suppose the mechanism on which an agent acts is disposed to respond to reasons. This is a property of the actual sequence, though one with implications for counterfactual situations. Consider an act that is caused by a tightly reason-responsive mechanism. Vivian is virtuous. If there had not been reason to do the act and Vivian had acted on the same mechanism as she actually acts on, she would not have done it. But in fact there is reason to do the act, and Vivian does it for that reason. A fortiori, she is able to do it. The ability to act in accord with reason does not here require an alternate sequence. Here we have both the ability to act in accord with reason and the operation of a reason-responsive mechanism. On both conditions, the act is responsible, and hence praiseworthy. Wolf's Reason View is

formulated in terms of the ability, not the responsive mechanism. But which one is really doing the work in supporting the intuitions to which her Reason View appeals?

To pry these conditions apart, we need to consider cases in which the agent does not do what reason demands. We have just seen that an implication runs from a responsive mechanism claim to an ability claim. From:

(1) Vivian acts on a mechanism that is tightly responsive to objective reasons

we get

(2) Vivian responds to objective reasons by doing the right thing

and thus:

(3) Vivian is able to do the right thing.

So, contraposing, if Vivian is unable to do the right thing, she cannot be acting on a mechanism that is tightly responsive to objective reasons.

However, the implication does not run in the other direction. Salome may be able to do the right thing, even though the mechanism on which she in fact acts is not tightly responsive to objective reasons. Perhaps it was possible outright for a different, more responsive mechanism to operate, so that it is true that Salome could have done the right thing for the reasons there were. But this is not the way things happened. Given the mechanism she actually acted on, she *would* not have done the right thing, even though she *could* have. The ability condition can be met even though the tight responsiveness condition is not met.

Consider cases in which neither the ability condition nor the responsive mechanism condition is met. If we judge someone is not responsible in such a case, is the lack of ability or the ill-disposed operative mechanism critical to our judgment? To find out, we can compare situations in which neither condition is met with cases in which the responsiveness

condition is not met but the ability condition is: cases in which someone would not have done the right thing, even though she could have. Does the sheer ability, the outright possibility on its own, make a difference to responsibility?

How is this strategy related to the strategy of Frankfurt cases? In a Frankfurt case the agent is not able to do otherwise because a fail-safe mechanism kicks in in the alternate sequence. But the operative causes are such that this mechanism never actually intervenes. The intuition Frankfurt cases appeal to is that this feature of the alternate sequence is irrelevant. An actual sequence condition sufficient for blame can be met, even if the alternate sequence could-have-done-otherwise condition is not.

This precise strategy for eliciting such an *irrelevant alternative intuition* cannot be pursued when the proposed actual sequence condition for blame is that someone is acting on a disposition that is tightly responsive to objective reasons. That is because of the inference from this particular tight disposition to ability. As we've seen in the case of Vivian, if someone is unable to do the right thing, she cannot be acting on a disposition that is tightly responsive to objective reasons. When the actual sequence condition is this tight, we cannot pry the actual sequence and the alternate sequence apart in the manner typical of Frankfurt cases, by supposing the actual sequence condition is met and the alternate sequence condition is not.¹⁸

However, Frankfurt cases are a symptom of something more general: the irrelevant alternative intuition. We can still elicit the irrelevant alternative intuition, *by prying this actual sequence disposition and the alternate sequence apart the other way round*. This is the key to the generalization of the irrelevant alternative intuition that emerges from comparison of Wolf and Fischer.

Suppose that an agent did not do the right thing and so did not act on a tightly reason-responsive mechanism but nevertheless was able to act rightly. She would not have done the right thing, even though she could have. That is, the mechanism on which she acted was not disposed to respond tightly to reasons, even though there was an alternate possible sequence in which another mechanism operated, leading to the right action. The question then is: what more can be said about the mechanism on which the agent did in fact act? Is it reason-responsive in some weaker way, that is compatible with being able but

failing to do the right thing? Could a plausible actual-sequence principle for blame be formulated in terms of such weaker reason-responsiveness? If so, we can then consider whether an agent is blameworthy because the actual sequence condition of weak reason-responsiveness is met, or because the agent is able to do the right thing. Given that she would not do the right thing even if she could, does it make any difference if in fact she can? Let's pursue this strategy for eliciting the irrelevant alternative intuition.

Consider the actual causes of action in cases in which an agent acts wrongly. These cannot be tightly responsive to objective reasons, or she would not act wrongly. But they might be more weakly reason-responsive: either loosely responsive, or responsive merely to subjective reasons. For example, the agent might act wrongly because she is weak-willed, or because she is evil. If she is *weak-willed*, her act may be caused in a way that is loosely but not tightly responsive to objective reasons: holding the mechanism on which she acts constant, she would respond to even stronger objective reasons to act in the right way, even though she does not respond to the objective reasons there actually are.¹⁹ If she is *evil*, her act may be caused in a way that is tightly responsive to evil subjective reasons, such as a desire to hurt others. These examples fall short of maximal reason-responsiveness: tight responsiveness to objective reasons. But they still illustrate weaker forms of reason-responsiveness: loose, or subjective.²⁰ A third way in which the agent could act wrongly, however, is by not acting on a mechanism that is even weakly reason-responsive. This might be the case, for example, for *medical* reasons: if the ordinary reason-responsive mechanisms that would otherwise operate are actually pre-empted by a brain tumor, or by mental illness, or by mad hypnotists or brain scientists.

We have three different dispositional characterizations of the mechanism on which an agent acts when she fails to do the right thing: as loosely reason-responsive, responsive only to subjective reasons, or not reason-responsive at all. These are all characterizations of the actual sequence. Each falls short in one way or another of maximal reason-responsiveness. In each case we can intuitively say that the agent would not have acted rightly, even if she could have.

We now need to bring into the picture alternate sequences and outright possibilities, what the agent could have done, all else constant, as opposed to what she would have done

under various counterfactual conditions, given the conditional dispositions of the operative mechanisms of action. Let's now compare two scenarios. In the first scenario, we stipulate that the agent in each of these three kinds of case *could* have acted rightly. In fact she acted on a mechanism such that she *would* not have acted rightly if she could have. Nevertheless, there was an alternate possible sequence, in which a different kind of mechanism operated. Since this kind of mechanism could have operated instead, the agent could have acted rightly. But as things actually were it played no causal role. In the second scenario, we stipulate instead that, unbeknownst to her, she could not have acted rightly. There was no alternate sequence. In neither case does the agent believe that she cannot act rightly. We now compare these two scenarios, and ask: If the agent would not have acted rightly even if she could have, does it matter whether or not she could have? If the actual causes of her act are not tightly responsive to objective reasons, so that she acts wrongly, and indeed are such that she would not act rightly even if she could, why is it relevant whether it is possible for her to act rightly? This possibility plays no explanatory role in what she actually does.

We can follow this recipe for each of our three actual-sequence characterizations, of cases of weakness of will, cases of evil, and medical cases. These questions pry apart various actual-sequence conditions from outright alternate sequence possibilities, and in particular the two conditions of weak reason-responsiveness from the outright ability to act rightly. Since Wolf appeals to the latter and Fischer to the former, this approach should zero in on what is at issue between their different reason-based accounts of responsibility.²¹

Consider first weak-willed Wilma. The causes of her action are such that she would not do the right thing for the reasons there are, even if she could, though she might do the right thing for stronger reasons, if there were any. Given that she would not do the right thing for the reasons there are, even if she could, does it make any difference to her responsibility if she could? Compare two cases. In one, Wilma would not have done the right thing for these reasons, even if she could have, and indeed she could have. In the other, Wilma would not have done the right thing for these reasons, even if she could have, and she couldn't have. In the first case but not the second there is an alternate sequence. In neither case does she believe she could not have done the right thing. Is Wilma any more

responsible in the second case? Is what matters for her responsibility the actual causes of her act and the dispositions they involve, or the outright possibility of an alternate sequence?

The descendant of Frankfurt's irrelevant alternative intuition here is that what matters for responsibility is not the outright ability to act rightly instead of wrongly, all else constant, but rather the operative disposition, whether one would act rightly if one could. When someone is weak-willed, her act is caused in such a way that she wouldn't have done the right thing for the reasons there were, even if she could have, yet she would have done the right thing if the reasons had been stronger. If this is enough for responsibility when there is an alternate possibility such that she could have done the right thing for the reasons there are, why not also when there is no such possibility, when she couldn't have done the right thing for the reasons there are? After all, as things are she wouldn't have done the right thing for the reasons there are, whether she could have or not. The outright possibility of an alternate sequence seems to be extrinsic, irrelevant, makes no difference, to her action or its causes.

These questions, then, express the irrelevant alternative intuition. She would have acted wrongly given the reasons there are, whether or not there was alternate possibility in which she did not act wrongly. So why should the alternate possibility matter to her responsibility? Why should the outright ability to act in accord with reason be necessary for responsibility if she wouldn't have acted that way even if she had been able to?

The irrelevant alternative intuition is conditional and comparative. It is not an intuition that Wilma is responsible for acting wrongly in either case. Rather, it is the intuition that *if* she is responsible for acting wrongly when she could have done otherwise, she is *no less* responsible when she couldn't have, given that she wouldn't have done otherwise whether she could have or not. The irrelevant alternative intuition ties the two cases together, but does not resolve whether we blame in both cases or neither. That turns on whether loose reason-responsiveness provides the right actual-sequence condition for blame, so that weak-willed acts are blameworthy in such cases. If we think that at least some weak-willed acts are be blameworthy, we need to loosen up a condition of tight reason-responsiveness in some way.

Recall the other two actual-sequence characterizations: subjective reason-responsiveness, and nonresponsiveness. Consider how the irrelevant alternative intuition could be elicited in these cases.

Ethel is evil. She desires to cause pain to others. We can suppose for the sake of argument that there are no objective reasons to do evil: no objective evil reasons. Nevertheless, suppose there can be subjective evil reasons, such as Ethel's desire to cause pain. That is, suppose that this desire, along with Ethel's belief that a certain act would be hurtful, do provide a subjective reason for her to do it. Now some people might act in ways that are tightly responsive to their subjective reasons, while others act in ways that are only loosely responsive to them. There is some interpretative play in the relations of subjective reasons to action; you can be weak-willed when your subjective reasons conflict. But let's suppose that the mechanism on which Ethel acts when she acts in accord with her subjective evil reasons is tightly responsive to these reasons. When she does the wrong thing, acting in accord with these evil reasons, it is true of her that she would not do as she does if she did not believe it to be hurtful. And when she does nothing, it is still true of her that if she believed she could do something hurtful, she would try to. Given that Ethel acts on this kind of mechanism when she acts in accord with her evil reasons, does it make any difference to her responsibility whether there is an alternate sequence of causes such that she is able to do the right thing, all else constant? Given that she would not do the right thing whether or not she could, why is it relevant whether she could? Either way, she actually does act on a mechanism that is tightly disposed to respond to evil subjective reasons. Even if she could act rightly, the alternate possibility would merely be standing by; it would do no work in explaining why she actually does what she does.

Again the irrelevant alternative intuition is conditional and comparative. *If* Ethel is responsible when she acts in accord with her evil desire and could have done rightly, then she is *no less* responsible when she acts in accord with the same evil reason but could not have done rightly, so long as it is true in both cases she *would* not have done rightly whether she could have or not. That ties the two cases together. The mere presence of the irrelevant alternative, a possibility that plays no role in explaining what she actually does, cannot make

a difference to her blameworthiness. It is a further question whether her disposition to evil makes her responsible in both cases, or in neither.

Finally, consider Maude. Maude has a medical problem. She also acts in a hurtful way in a particular case, but the mechanism on which she acts is not reason-responsive in general, either loosely or to subjective reasons. Her action is a symptom of her mental illness, and is caused by abnormal neurotransmitter fluctuations that bear no projectible relationship to reasons. On a reason-based view of responsibility, such a causal history would tend to defeat responsibility for an act. If the mechanism on which Maude acts is not responsive to reasons at all and as a result as things are she would not do the right thing even if she could, then an actual-sequence reason-based view of responsibility will find her not responsible. Does it make any difference whether she indeed could have done the right thing? Suppose there is an alternate sequence of causes in which a different, reason-responsive mechanism operates, such that there is an outright possibility that she does the right thing. Does this make Maude any more responsible? After all, this alternate sequence plays no role in the actual causes of her act. The irrelevant alternative intuition is again conditional and comparative: the outright ability to do the right thing makes Maude no *more* responsible, given that in fact she acts on a nonresponsive mechanism.

5. More on how the Irrelevant Alternative Intuition Generalizes the Point of Frankfurt Cases.

At this point we can stand back and consider where our exploration of reason-based views of responsibility has led. Wolf's Reason View is able to ignore the distinction between actual and alternate sequence principles, as well as Frankfurt cases involving blame, because the link she requires to reason is a tight link to objective reasons. This is what yields her asymmetry between praise and blame with respect to alternate possibilities.

Moreover, an actual sequence condition of tight objective reason-responsiveness makes responsible weakness of will or evil-doing impossible.²² Fischer is right that weakness of will can in principle be responsible (1994, 168), and surely the same is true of evil-doing. A reason-responsiveness principle of responsibility that rules out these possibilities in principle cannot be correct. Thus, a reason-responsiveness condition of

responsibility should be less-than-maximal, so as to allow these possibilities in principle. Such a principle would not support the asymmetry between praise and blame.

We can then distinguish the responsiveness to reasons of the mechanism on which the agent acts from the agent's outright ability to act rightly. This distinction allows us to set up cases in which it is supposed that the agent *wouldn't* have done otherwise or have avoided blame, even if she *could* have. Making sense of this intuitive supposition is critical to eliciting the irrelevant alternative intuition. The irrelevant alternative intuition then emerges: the actual sequence of operative causes is where the action is. That is what responsibility turns on. If the actual sequence is such as to support or defeat responsibility, adding or subtracting alternate sequences makes no difference.²³

The irrelevant alternative intuition generalizes the point of Frankfurt cases in a significant way. Those cases involve a blameworthy actual sequence and a blame-avoiding alternate sequence, and so do our cases. In Frankfurt cases, the agent couldn't have done otherwise because in the alternate sequence she does the same thing as she does in the actual sequence. The alternate sequence in those cases is associated with a counterfactual intervener, a fail-safe mechanism which guarantees that in the alternate sequence the agent does not do otherwise. The intuition in Frankfurt cases is that what matters is the way the action is actually brought about, not whether the agent could have done otherwise. By contrast, in our cases no fail-safe mechanism guarantees that in the alternate sequence the agent does not do otherwise. But that guarantee turns out not be critical; the intuition survives that what matters is the way the action is actually brought about, not whether the agent could have done otherwise. What the intuition does depend on is making sense of the supposition that the agent would not have done otherwise, whether or not she could have done otherwise.²⁴

So far, this supposition has been understood by reference to Fischer's mechanism-based account of responsibility. However, there may be other ways to explain it. For example: "*A wouldn't* have done otherwise, whether or not he *could* have" might be understood as follows. Suppose *A* actually does *X*, where doing *X* is acting in a weak-willed way. Suppose also that the actual world is deterministic, so that there is no alternate possibility to *A*'s doing *X*. Then there's a set of possible indeterministic worlds different from

the actual world only in that these possible worlds contain two alternate possibilities: in each of these possible worlds, unlike the actual world, it is possible that *A* does *X* and also possible that *A* does otherwise, so does not act in a weak-willed way. Call these indeterministic worlds the “alternate possibility worlds”, or AP worlds. In different subsets of the set of AP worlds, different alternate possibilities are actualized. There is a subset of the AP worlds in which *A* still acts out of weak will and does *X*, and another subset in which *A* does not act out of weak will and does indeed do otherwise. To claim that *A* wouldn’t have done otherwise whether or not he could have, or that *A* would have acted out of weak will whether or not he could have acted otherwise, would be to claim that the AP worlds in which *A* acts out of weak will are closer to the actual deterministic world than the AP worlds in which *A* does otherwise.

What could make this claim true? The AP worlds in which *A* acts out of weak will could be closer to the actual deterministic world because they are similar in respect of *A*’s weak-willed disposition, of which there might be both deterministic and indeterministic realizations. A comparison between weakness of will and control may help to make the point. There could be deterministic or indeterministic realizations of a given control system software (see Clarke 1995). Neither realization may be perfect, and there may be trade-offs between the kinds of faults generated by deterministic and indeterministic hardware which leave the two realizations functionally equivalent. Evolution, for example, might have used either equally well; natural selection might have operated on either in a similar way. In a closely related sense, there could be deterministic or indeterministic realizations of a weak-willed disposition, which are functionally equivalent.

Compare this account to Frankfurt’s suggestion that the fact that an individual couldn’t have done otherwise is relevant to responsibility only if he did what he did *only because he could not have done otherwise* (1969, 837-839). What is it to do what you do only because you cannot do otherwise? Whatever it is, it must be the case that you cannot do otherwise and that this plays an essential role in why you do what you do.

Frankfurt says that if someone could not have done otherwise but did not do what he did because he could not have done otherwise, then he would have done the same thing even if he could have done otherwise. The circumstances that made it impossible could

have been subtracted from the situation without affecting what happened or why. Whatever made someone do what he did would still have done so, even if he could have done otherwise (1969, 837). In Frankfurt cases, the circumstances that make it impossible to do otherwise are in the alternate sequence: the counterfactual intervenor. These could be subtracted without affecting what happened or why because they are not what made the person do what he did: he acted for his own reasons. But what if the circumstances that make it impossible to do otherwise are in the actual sequence, in a deterministic world? The comparison shows yet again how the irrelevant alternative intuition generalizes the point of Frankfurt cases. To spell this out:

If the world is indeterministic but you are in a Frankfurt case, you could not have done otherwise. In such a case, this is because in the alternate world, the counterfactual intervenor would make you do the same thing as you actually do. You do what you do for your own reasons, and you could not have done otherwise because of the counterfactual intervenor. Do you do what you do only because you could not have done otherwise? No; since even if the counterfactual intervenor had not been there, you would have done what you did, for your own reasons.

If the world is deterministic, then you cannot do otherwise, because of the past and the laws. Indeed, you do what you do, because of the past and the laws. But do you do what you do only because you cannot do otherwise? Not necessarily. You might have done the same even if you could have done otherwise, that is, even if the world were not deterministic. At this point, my account of what it could mean to say that someone “wouldn’t have done otherwise, even if she could have” may be relevant to explaining what Frankfurt’s “only because she could not have done otherwise” formulation amounts to. For example, if you would have acted out of weakness of will even if you could have done otherwise, in the sense that the AP worlds in which you are weak-willed are closer to the actual deterministic world than the other AP worlds, then you didn’t do what you did only because you couldn’t have done otherwise. Closeness to the actual world might reflect, for example, the functional or evolutionary equivalence of deterministic and indeterministic realizations of a disposition.

This suggestion generalizes Frankfurt's suggestion, in that it does not depend on a counterfactual intervenor to make sense of the supposition that someone cannot do otherwise. We can start with a situation, as in my weakness of will case, in which someone cannot do otherwise not because she is in a Frankfurt case, but simply because she is in a deterministic world. So my suggestion frees us from the somewhat peculiar context of Frankfurt cases with their fail-safe mechanisms, and returns us to the more general deterministic context in which could-have-done-otherwise worries have traditionally arisen. It thus shows that the importance of the basic irrelevant alternative intuition behind Frankfurt cases cannot be avoided by objections to the peculiarity of Frankfurt cases, with their counterfactual intervenors or other sci-fi fail-safe apparatus.

In sum, the question whether you would have done otherwise whether or not you could have is closely related to the question whether you did what you did only because you couldn't have done otherwise, in the ways just indicated. My suggestion can thus be regarded as a way of developing and generalizing the view that what matters is whether you did what you did only because you couldn't have done otherwise.

Yet one further illustration of the way in which the irrelevant alternative intuition generalizes Frankfurt's original point may be helpful. Michael Otsuka has recently argued (1998) that a principle that requires for blameworthiness that the agent could have avoided blame is resistant to the Frankfurt-style counterexamples that embarrass a could-have-done-otherwise principle. His basic point turns on the fact that in a Frankfurt case, even though the agent could not have done otherwise than perform an act of a given type, the agent could have avoided blame for performing an act of that type. If Jones had wavered and induced Black's fail-safe mechanism to take over, then Jones would not have been blameworthy. Intuitively, Jones is not exempt from responsibility in such a case, even though he could not have done otherwise. Otsuka traces this intuition to Jones' satisfaction of a different alternate sequence condition, namely, to the fact that he could have avoided blame, which Otsuka recommends. I here show that the Irrelevant Alternative Intuition operates against the Otsuka's could-have-avoided-blame principle just as effectively as it does against the could-have-done-otherwise principle.

Compare two cases. The first, indeterministic situation is a Frankfurt case, where there are just two possible sequences. In the alternate ‘waver’ sequence in which the counterfactual intervention occurs, the agent would avoid blame for her bad, weak-willed act. If the mad scientist had predicted she would waver, he would have ensured that she acted in the same way. But he did not. The counterfactual intervention is merely possible; it never actually occurs, and the fail-safe device never actually operates. In fact the agent does the act for her own reasons, without intervention. Still, she could have avoided blame because the alternate waver sequence could have occurred instead. If it had, and the fail-safe device had operated, she would not have been blameworthy. Otsuka suggests that an agent is intuitively responsible in this kind of case because she could have avoided blame, even if though she could not have done otherwise.

In the second, deterministic case, there is only one possible sequence, no alternate ‘waver’ sequence and no possible counterfactual intervention. So the agent could not have avoided blame. Nevertheless, in both cases she acts in the same way, on the same kind of disposition and for the same reasons. In both cases, it is true of her that she would not have avoided blame even *if* she could have.²⁵ That is, in both cases, the dispositions involved in the processes by which the action is brought about are such that she would not have avoided blame, even *if* she could have. But in one case she could have, and in the other case she could not have; in one case but not the other the “if she could have” condition is counterfactual. The only difference between the cases is that in the indeterministic case there is an alternate sequence, an outright possibility of waver and counterfactual intervention, and there is no such possibility in the other, deterministic case. That is, in the indeterministic case there is an outright causal possibility that the relevant action might have been brought about in a different way, by a functionally different kind of process, holding everything else constant. But in neither case was the action actually brought about in this different way, by this different kind of process. The difference is one in what might have happened, all else constant. It is not a difference in what did happen or in how what did happen was caused. Nor is it a difference in the dispositional properties of the process by which the action was actually brought about.

Is the agent any more blameworthy when in addition she could have avoided blame? Or is what matters the fact that even *if* she could have, she wouldn't have? Does the existence of the alternate possibility matter, or the dispositional character of the actual causes of the act?

The descendent of the Frankfurt intuition here is that what matters is the dispositional character of the actual causes of her act, the supposition that she would not have avoided blame even if she could have. Given that, the alternate sequence is an *irrelevant alternative*, an extrinsic difference between the cases that does not affect her responsibility. The outright possibility of avoiding blame is not what matters for responsibility, on this intuition, but rather the dispositional character of the way in which the action is actually brought about.

The outright possibility of an alternate sequence has traditionally been regarded as necessary for responsibility. By contrast, the irrelevant alternative intuition casts it in an extraneous role, as a mere piece of luck, not a feature of the agent that is operative in the right way to be relevant to her responsibility. If the way in which the action is brought about is such that the agent would not avoid blame even if she were able to, but would act on her own bad reasons whether or not she could avoid blame, then the features of her and of her act that are relevant to blame are the same in the two cases. Why should it matter to her blameworthiness whether an alternate possibility exists? Her victims may think "It didn't have to happen!" and that may make their plight harder to bear. But she is just as bad and her act just as intrinsically blameworthy, just as much a proper object of resentment, either way. The intuition is that responsibility is independent of such irrelevant alternatives.²⁶

6. A Remaining Issue: Reason-responsiveness vs. Regression.

The main purpose of this article has been to develop and explore the irrelevant alternative intuition in relation to Wolf's and Fischer's reason-based accounts of responsibility. However, this intuition also prompts us to look harder at the character of the actual sequence, and at a regression requirement in particular. Someone might agree that two cases in a Frankfurt-type comparison do not differ in blameworthiness in virtue of

the alternate sequence, but think that we need to know more about the actual sequence to know whether the agent is blameworthy in both cases or in neither. Suppose that in neither case, and whether or not determinism holds, is the agent regressively responsible for the causes of her act, all the way back. But suppose that in both cases the action is brought about in a way that is responsive to evil subjective reasons, such as a desire to cause pain to others.²⁷ When her action is brought about in this way, she is disposed to act in the way she does just when doing so will hurt her victim, in order to hurt him. Of course, on the waver scenario, her action would not be brought about in the same way, but would rather be brought about in a different way, involving the counterfactual intervention. But in fact intervention does not occur, and her action is actually brought about in this evil-responsive way. Is she blameworthy in both cases or in neither?

These suppositions pit two actual sequence principles of responsibility against one another: a regression requirement, and a requirement that the act be brought about in a way that is reason-responsive (Fischer 1994, 1998). The latter does not require responsibility all the way back along the actual sequence, but rather imposes a dispositional condition locally on the causes of the act. When regression is pitted against reason-responsiveness, does intuition favor the latter? These issues are not pursued here, but a regression requirement remains a major source of challenge to reason-responsiveness accounts of responsibility.

Having traced through the issues between Wolf's version and Fischer's (1994) version of a reason-based view of responsibility, we can see each of them as making good gaps in the other's discussion. Wolf fails to consider the implications of Frankfurt cases for reason-based views where blame is involved, which Fischer addresses. However, Fischer (1994) fails to consider the challenge to a reason-based view of responsibility from a regressive conception of responsibility, which in turn Wolf addresses (and which gets more attention in Fischer and Ravizza 1998, ch. 7, 8). The challenge is: how other than in terms of the failure of regression can we explain why responsibility is intuitively lacking in various cases? We could respond by making a comparison within the class of actual-sequence principles of responsibility: between a less-than-maximal reason-responsiveness principle and a regression principle. As just indicated, we would need to consider cases in which an actual-

sequence reason-responsiveness principle is met but the regression principle is not, or vice versa.

But while a regression requirement may threaten reason-responsiveness accounts of responsibility, these rivals are all in the category of actual-sequence approaches to responsibility. As an actual-sequence requirement, a regression principle does not directly support an alternate sequence principle such as could-have-done-otherwise. Since the irrelevant alternative intuition cuts against alternate-sequence approaches, the intuition is not directly threatened by a regression requirement.²⁸

References

Adams, Robert Merrihew (1985) "Involuntary Sins", *Philosophical Review* XCIV (1), 3-31.

Clarke, Randolph (1995) "Indeterminism and Control," *American Philosophical Quarterly* 32 (2), 125-37.

Fischer, John Martin (1982) "Responsibility and Control," *The Journal of Philosophy* LXXIX (1), 24-40.

--- (1983) "Incompatibilism," *Philosophical Studies* 43, 127-37.

--- (1987), "Responsiveness and Moral Responsibility," in *Responsibility, Character, and the Emotions*, Ferdinand Schoeman, ed. Cambridge, UK: Cambridge University Press, 81-106.

--- (1994), *The Metaphysics of Free Will: An Essay on Control*. Oxford, England: Blackwell.

Fischer, John Martin and Mark Ravizza (1991) "Responsibility and Inevitability," *Ethics* 101, 258-78.

--- (1992) "The Inevitable," *Australasian Journal of Philosophy* 70, 388-404.

Fischer, John Martin and Mark Ravizza (1998), *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge, England: Cambridge University Press.

Frankfurt, Harry G. (1969) "Alternate Possibilities and Moral Responsibility," *The Journal of Philosophy* LXVI (23).

--- (1971) "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy* LXVII (1), 5-20.

--- (1973), "Coercion and Moral Responsibility," in *Essays on Freedom of Action*, Ted Honderich, ed. London: Routledge & Kegan Paul, 65-86.

--- (1976), "Identification and Externality," in *The Identities of Persons*, Amelie Rorty, ed. Berkeley, USA: University of California Press, .

--- (1987), "Identification and Wholeheartedness," in *Responsibility, Character, and the Emotions*, Ferdinand Schoeman, ed. Cambridge, UK: Cambridge University Press, 27-45.

Hurley, S. L. (1989), *Natural Reasons: Personality and Polity*. New York: Oxford University Press.

Honore, Tony (1988), "Responsibility and Luck", *Law Review Quarterly*, 530-553.

Klein, Martha (1990), *Determinism, Blameworthiness and Deprivation*. Oxford, England: Clarendon Press.

Otsuka, Michael (1998) "Incompatibilism and the Avoidability of Blame", *Ethics* 108, 685-701.

Powers, William T. (1973), *Behavior: The Control of Perception*. Chicago: Aldine.

Shatz, David (1986), "Free Will and the Structure of Motivation", *Midwest Studies in Philosophy X*, 451-482.

Wallace, R. Jay (1996), *Responsibility and the Moral Sentiments*. Cambridge: Harvard University Press.

Wolf, Susan (1980) "Asymmetrical Freedom," *The Journal of Philosophy* LXXVII (3), 151-66.

--- (1987), "Sanity and the Metaphysics of Responsibility," in *Responsibility, Character, and the Emotions*, Ferdinand Schoeman, ed. Cambridge, UK: Cambridge University Press, 46-62.

--- (1990), *Freedom Within Reason*. New York, USA: Oxford University Press.

NOTES

I am very grateful to the Nuffield Foundation for the award of a Social Science Research Fellowship that gave me the freedom from other duties to pursue this research (as part of a larger project). In addition, for helpful comments on earlier drafts of this material and discussion of related ideas, I would like to thank Karin Boxer, Gerald Cohen, Mark Greenberg, Martha Klein, Michael Otsuka, Derek Parfit, Andrew Williams, Christopher Woodard, and several anonymous referees.

¹ And perhaps even further back, to Aristotle.

² See Shatz 1986, sect. VII, for a discussion that locates contemporary reason-sensitivity accounts of responsibility in the broader context of contemporary approaches to responsibility.

³ Klein 1990 also draws this distinction.

⁴ See, for example, Klein 1990.

⁵ At that points Wolf's criticisms of the Real Self view *seem* to commit her to a regression requirement. But since Wolf does not ultimately accept a regression requirement, her criticisms of the Real Self view should be either recast or reinterpreted so that they clearly do not depend on a regression requirement.

⁶ This assumption might be unwarranted if objective reasons for acting can come into ineliminable conflict with each other, so that a correct view of how to resolve the conflict does not eliminate the conflict or subsume the force of the overridden reason without remainder. (I argue that this is possible in Hurley 1989, ch. 7.) Where there is such a conflict, someone might be weak willed and do the wrong thing, all things considered, but nevertheless act for the conflicting objective reason.

⁷ Although, as indicated, her view in effect expresses an asymmetrical form of the irrelevant alternative intuition, which I claim is also at work in Frankfurt cases.

⁸ In Fischer and Ravizza 1998, Fischer revises his 1994 view in various ways. On the revised view, weak reason-responsiveness is too weak a condition for blameworthiness. See note 16 below.

⁹ This and other varieties of reason-responsiveness are explained further in the following section; see Fischer 1994, 244n; Fischer and Ravizza 1991, 277; Fischer and Ravizza 1998.

¹⁰ 1994, 151, 205; though Fischer appears to have changed his mind about this in Fischer and Ravizza 1998.

¹¹ See Klein 1990, who argues this, on the assumption that we are not responsible for the causes of our acts. The regression requirement can be regarded as a generalization of Klein's U-condition, though she does not use the term "regressive".

¹² "Even an agent who acts against good reasons can be responsive to some reasons" (Fischer 1994, 167). Is this because the reasons in question need not be objective, or because the responsiveness in question needs only to be loose? Does pluralism about reasons allow that there are bad but objective reasons? In one example, Sam's reasons for killing are bad reasons, but "they are his reasons", and Sam is viewed as responsible and blameworthy (Fischer and Ravizza 1991, 258). If the counterfactual intervenor tracked objective reasons tightly, but these were not the agent's subjective reasons, would the alternate-sequence mechanism count as reason-responsive? For more recent discussion, see Fischer and Ravizza 1998. Also relevant here are the questions raised in note 7 above: can objective reasons for action come into ineliminable conflict? Can it be the case that when someone acts in a weak-willed way against his correct better judgment he nevertheless acts on and for an objective reason?

¹³ However, Fischer and Ravizza 1991 did endorse a different asymmetry, between responsibility for acts and for omissions. On their past view, responsibility for an act does not require the freedom to refrain from performing the act, while responsibility for failure to perform an act does require the freedom to perform (1991, 271; cf. Fischer 1994). Their arguments for this view seem to rest on a misunderstanding of control and its relationships to reason-responsiveness and to alternative possibilities, and to miss the point of control theory. They have recently revised this element of their view (Fischer and Ravizza 1998).

Their earlier position was that since omissions do not involve "actual causal control" of movement, but responsibility does require reason-responsive control in some form, responsible omissions must instead involve "regulative control". The latter is an alternate sequence requirement, demanding outright freedom to perform. The misunderstanding can be corrected in Fischer's own terms.

Reason-responsive control involves dispositional properties of operative mechanisms in the actual sequence of events. It requires that a different response would be made under different conditions, for example, if the reasons were different or if the circumstances different and so

differently related to the same reasons. But it does not require that a different response could be made, holding all else constant. It does not require that an alternate sequence be possible, outright. These points apply to omissions as much as to acts. If the operative mechanisms when you omit to act are so disposed that you would have acted under different conditions and these dispositions characterize a reason-responsive control system, then your omission may nevertheless count as reason-responsive and as responsible. For example, you may be responsible for omitting something you should have done in virtue of reason-responsive dispositions characteristic of weak will, or of dedication to evil. For further discussion, see sects. 3 and 4 below, and note 21.

Fischer and Ravizza 1998 no longer require for responsibility for omissions the kind of control that involves alternate possibilities. However, they still think that there is a kind of control, indeed the kind typically associated with moral responsibility, which they call “regulative control”, that requires alternate possibilities (20, 24, 37, 338). By contrast, I hold that *no* kind of control *requires* the outright possibility of an alternate sequence, all else constant (see also Wallace 1996, 86-87, 181, 189-90, 220-21, 262-63, who argues for powers of reflective self-control as conditions of accountability and that such powers are compatible with determinism). Although control is compatible with indeterminism, ‘regulative control’ is not a kind of control at all. Moreover, Fischer and Ravizza still in 1998 find it natural to think of the reason-responsiveness of an agent as requiring alternate possibilities (1998, 37), even if the reason-responsiveness of a mechanism does not. By contrast, I hold that agent reason-responsiveness no more requires alternate possibilities than mechanism reason-responsiveness, for the reasons given in the preceding paragraph.

A different and more promising asymmetry, which Fischer’s view does support, is one between talents and handicaps with respect to reason-responsiveness. Omissions to act in worthwhile ways are often explained not by the operation of mechanisms that are reason-responsive in these blame-supporting ways, but rather by sheer lack of reason-responsiveness. No suggestion is intended that people with various independently identified handicaps are irrational, though some handicaps may compromise rationality. Rather, the suggestion is the other way around: sheer lack of reason-responsiveness can itself be regarded as a kind of handicap, and reason-responsiveness as a generic talent. People handicapped by lack of reason-responsiveness may make poor quality decisions, both to act and to omit, for which they are not responsible, even though talented, reason-responsive types are responsible for their acts. This point may be relevant to the roles of responsibility in theories of justice.

¹⁴ He uses the terms “strong” vs. “weak” reason-responsiveness (1994, 164ff).

¹⁵ See Fischer 1994, 165-167. As explained in the text, Wolf’s view is not a pure actual sequence, reason-responsiveness view; her ability condition makes alternate sequence requirements. She thinks that it is possible to be able to act in accord with the True and the Good, but fail to do so nonetheless, thus that weakness of will is possible (1990, 88-89).

¹⁶ It might be thought that loose reason-responsiveness is too weak a condition for blameworthiness but that tight reason-responsiveness is too strong (Fischer and Ravizza move to this view in 1998). What significant middle ground might there be between responding to reason in all possible worlds and only in some possible worlds? Can some subset of possible worlds be specified that yields a condition intermediate between loose and tight reason-responsiveness? Further consideration of weakness of will suggests one way to do this (cf. Fischer and Ravizza 1998, ch. 3).

Distinguish all-things-considered reasons to act (“It’s the right thing to do”) from specific reasons to act (“That would be the kindest thing to do, though not necessarily the fairest”). Assume that the latter may conflict, even if they are objective reasons (see Hurley 1989, ch. 7, 8 on *pro tanto* reasons). A weak-willed agent does not act in accord with the reasons she has, all things considered, but she does act for one of the conflicting specific reasons she has. Suppose distinct specific reasons conflict over whether to do *x* or *y*, the agent deliberates and judges correctly that all-things-considered the right thing to do is *x*, and instead she does *y*. Moreover, she does *y* for the specific reason there is to do it, even though her all-things-considered deliberated judgement to the contrary took account of this reason. In this familiar type of situation, intermediate reason-responsiveness might require the mechanism on which the agent acts not just to respond to reason in some other possible worlds, but to be responsive to deliberation in the face of conflicting reasons in some other possible worlds. More specifically, it could require the mechanism on which the agent acts to satisfy various combinations of the following conditions: (a) that it responds to specific reasons (of the types in play in the actual world) in some other worlds where reasons of those types obtain; (b) that it responds to all-things-considered, deliberative reason in some other worlds where specific reasons conflict; (c) that it responds to all-things-considered, deliberative reasons in some other worlds when specific reasons of the types in play in the actual world obtain and also conflict. Each of these conditions could be strengthened to require ‘most other worlds’ rather than ‘some other worlds’. (See and cf. Honore: “...a person guilty of fault must have, besides a general capacity for decision and

action, the ability to succeed most of the time in doing the sort of thing which would on this occasion have averted the harm.” 1988, 531; see also 550-551.)

If we recognize intermediate kinds of reason-responsiveness between loose and tight responsiveness, we get another two categories of responsiveness: intermediate responsiveness could be required merely to subjective reasons, or to objective reasons. Some form of intermediate reason-responsiveness may be the most attractive kind of reason-responsiveness for purposes of formulating principles of responsibility. See Fischer and Ravizza 1998 on what they call “moderate reason-responsiveness”. However, intermediate reason-responsiveness raises complications beyond the scope of this article, which the article’s point does not depend on. So it will not be developed further here; we can still learn something from the other possibilities.

¹⁷ Fischer is sensitive to worries about how mechanisms are individuated, and to parallel worries in epistemology (1987, 1994, Fischer and Ravizza 1998). For present purposes we simply help ourselves to his use of the idea of the dispositional character of the mechanism that actually operates when an agent acts. See discussion in sect. 4.

¹⁸ This is the kernel of truth in Wolf’s asymmetry thesis. However, it turns critically on the tightness of the reason-responsive disposition; a looser formulation will miss Wolf’s point (as in Fischer’s weak reason-responsiveness; cf. Fischer and Ravizza 1991; Fischer 1994, 156ff).

¹⁹ Perhaps she also satisfies an intermediate requirement of responsiveness: she may be generally responsive to the results of rational deliberation about what should be done, all things considered, when objective reasons conflict.

²⁰ Someone could be both evil and weak-willed, but that doesn’t raise any different issues. Reason-responsiveness that is both loose and subjective is not considered separately.

²¹ It may be tempting to ask at this point: when we stipulate that a different mechanism could have operated but didn’t, was whether it operated or not under the agent’s control? But this confuses an alternate sequence condition with an actual sequence control condition. Control of *X* requires that *X* would have been otherwise under certain conditions, not that it could have been otherwise all else constant. Whether a different mechanism could have operated is a question about an alternate sequence, about outright possibility, all else constant. But control, like reason-responsiveness, involves dispositional properties of the actual sequence. If the agent controls which mechanism operates, then the agent maintains the state of affairs with respect to mechanism-operation at some target value in the face of

exogenous disturbances, by adjusting her behavior accordingly to compensate for those disturbances (see for example Powers 1973). Control depends on dispositions to respond so as to maintain the target value under various counterfactual conditions. Moreover, if you are controlling something, you only do otherwise if circumstances are different: different exogenous disturbances demand different responses to maintain the target value of what you control. (Fischer and Ravizza 1998 reinvent some of the wheels of control theory, e.g. at 120.) If your variation in behavior is independent of those conditions, you are not maintaining the target value, hence not controlling whatever takes that value; target maintenance is essential to control. (Hence what Fischer and Ravizza 1998 still call ‘regulative control’ is not a kind of control at all.) You may in turn control the target value, but that just pushes the conditional-dispositional character of control back a step and begins a control hierarchy. And similar points apply to control of the target value and to control hierarchies. Meta-control no more requires alternate possibilities than does simple control (cf. Fischer and Ravizza 1998, 31).

We are trying to evaluate whether an alternate sequence condition of responsibility holds in addition to an actual sequence reason-responsiveness condition. To do this we need to separate actual sequence conditions from the outright possibility of an alternate sequence, not run them together. I discuss the dispositional character of control more fully elsewhere. But for now note that a control requirement is closer to a requirement of reason-responsiveness than to one of outright alternate possibility, all else constant. (These points are elaborated in *Justice, Luck, and Knowledge*, in progress.)

These remarks about the dispositional character of control can easily be adjusted to allow that control is compatible with indeterministic causation (Clarke 1995). Control is neutral with respect to determinism. Cf. Otsuka 1998, part I.

²² But, again, since Wolf’s view is not a pure actual sequence view this remark does not strictly apply to her. See note 15 above.

²³ Is this just a new way of elaborating the old intuition that indeterminism cannot make for responsibility?

²⁴ For the role of this supposition in Frankfurt's work, see e.g. 1991, 19; 1969, 837. This article can be regarded as emphasizing and trying to explain further those few passages in Frankfurt in which the "wouldn't have done otherwise even if could have" formulation appears.

²⁵ Causal mechanisms and dispositions can make counterfactuals true in both deterministic and indeterministic worlds (see Clark 1995 for related points).

²⁶ This phrase suggests an analogy to intuitions supporting independence or sure-thing axioms in rational choice, which would be interesting to pursue but which is not developed here.

²⁷ Since it is plausible that agents can be more or less tightly disposed to respond to their own desires or subjective reasons as well as to objective reasons, the mechanism on which an agent acts should not be thought of simply as his desires. Rather, the mechanism on which the agent acts may be disposed to respond to various reasons, which may include desires, more or less tightly. It is possible for reason-responsive mechanisms on which agents act to mediate causally between reasons and actions in various ways. Responsiveness to subjective reasons and other types of reason-responsiveness are discussed further in sections 3 and 4 below.

²⁸ See Klein 1990 on the distinction between regression and could-have-done-otherwise principles; *Justice, Luck, and Knowledge* (in progress) argues that regression and could-have-done-otherwise requirements are mutually independent.