This article was downloaded by: *[University of Notre Dame]*
On: *24 February 2011*
Access details: *Access Details: [subscription number 917395010]*
Publisher *Routledge*
Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK

Australasian Journal of Philosophy
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713659165

The inevitable

John Martin Fischer[a]; Mark Ravizza[a]
[a] University of California, Riverside

To cite this Article Fischer, John Martin and Ravizza, Mark(1992) 'The inevitable', Australasian Journal of Philosophy, 70: 4, 388 — 404
To link to this Article: DOI: 10.1080/00048409212345271
URL: http://dx.doi.org/10.1080/00048409212345271

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.informaworld.com/terms-and-conditions-of-access.pdf

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Australasian Journal of Philosophy
Vol. 70, No. 4; December 1992

THE INEVITABLE

John Martin Fischer and Mark Ravizza

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

In this paper we wish to explore some issues pertinent to moral responsibility for the consequences of human actions. We shall begin by thinking about some puzzling examples and the approaches for analyzing those examples suggested by certain philosophers.[1] We shall criticize and reject these approaches, and we shall sketch our own alternative theory of moral responsibility for consequences.[2] Although our presentation is only a sketch of a theory, we show how our approach analyzes and systematizes the relevant examples in a perspicuous fashion. This provides some reason to think that it will be worthwhile further to develop and refine the theory.

In an interesting paper, 'Incompatibilism Without the Principle of Alternative Possibilities', Robert Heinaman lays out an account of moral responsibility for consequence-universals; that is, (among other things) he develops an account of the conditions for moral responsibility for the consequences of what one does, where these consequences are construed as universals rather than particulars.[3] Having set out this account, Heinaman defends it against certain objections which he constructs on behalf of Peter van Inwagen.[4] In arguing against van Inwagen, Heinaman relies on principles (I) and (II):

> (I) If S is responsible for a state of affairs A, then if 'F' is any false statement (other than 'not-A'), then S is responsible for the state of affairs that A or F.
> (II) If S is not responsible for A, then for any proposition P, S is not responsible for the state of affairs that A or P.

1 Here we focus on a few salient alternative approaches. For a more comprehensive discussion of the full range of theories of moral responsibility for consequences, see J.M. Fischer and M. Ravizza, 'Responsibility for Consequences' in J. Coleman and A. Buchanan (eds), *In Harm's Way: Essays in Honour of Joel Feinberg* (Cambridge: Cambridge University Press, forthcoming 1993).

2 We have presented versions of this theory in J.M. Fischer, 'Responsiveness and Moral Responsibility' in F. Schoeman (ed.), *Responsibility, Character, and the Emotions* (Cambridge: Cambridge University Press, 1987), pp. 81-106; and J.M. Fischer and M. Ravizza, 'Responsibility and Inevitability', *Ethics* 101 (1991) pp. 258-278; and 1992 (forthcoming).

3 R. Heinaman, 'Incompatibilism Without the Principle of Alternative Possibilities', *Australasian Journal of Philosophy* 64 (1986) pp. 266-276. Heinaman's discussion of responsibility for consequence-universals is one strand among various interesting strands of discussion.

4 P. van Inwagen, 'Ability and Responsibility', *The Philosophical Review* 87 (1978) pp. 201-224. van Inwagen is a proponent of the traditional view that associates moral responsibility with alternative possibilities; specifically, van Inwagen holds that in order to be morally responsible for a consequence of what one does, one must have been able (at some relevant time) to *prevent* that consequence from obtaining.

As a springboard for our discussion, we shall criticize Heinaman's account of moral responsibility for consequence-universals. Our critique of Heinaman's proposal will point also to the inadequacy of Heinaman's principle (I). Further, we shall argue that (II) is false. Thus, Heinaman's defence of his account against van Inwagen-type objections is not cogent. But we also argue that van Inwagen's approach to moral responsibility for consequences is inadequate. Finally, we develop our own account of moral responsibility for consequence-universals and explain how it systematizes the relevant examples in a way which renders it superior to *both* the approaches of Heinaman and van Inwagen.

I. Heinaman's Argument

Heinaman wishes to say that there are states of affairs which are inevitable for an agent but for which the agent is nevertheless morally responsible. We agree. If an agent freely brings about a state of affairs, then he may be held responsible for that state of affairs even though he could not have prevented it obtaining. Here is an example of this sort.[5]

Sam confides in his friend, Jack. Sam tells Jack of his plan to murder the mayor of the town in which they live. Sam is disturbed about the mayor's liberal policies, especially his progressive taxation scheme. Whereas Sam's reasons for proposing to kill the mayor are bad ones, they are *his* reasons: he has not been hypnotized, brainwashed, duped, coerced, and so forth. Sam has deliberated coolly, and he has settled on his murderous course of action.

Sam is bad, and Jack is no better. Jack is pleased with Sam's plan, but Jack is a rather anxious person. Because Jack worries that Sam might waver, Jack has secretly installed a device in Sam's brain which allows him to monitor all of Sam's brain activity and to intervene in it, if he desires. The device can be employed by Jack to insure that Sam decides to kill the mayor and that he acts on this decision; the device works by electronic stimulation of the brain. Let us imagine that Jack is absolutely committed to activating the device to insure that Sam kills the mayor, should Sam show any sign of not carrying out his original plan. Also, we can imagine that there is nothing that Sam could do to prevent the device from being fully effective, if it is employed by Jack in order to cause Sam to kill the mayor.

Sam and Jack both go to a meeting at the town hall, and Sam methodically carries out his plan to kill the mayor. He does not waver in any way, and he shoots the mayor as a result of his original deliberations and plan. Jack thus plays absolutely no role in Sam's decision and action; the electronic device simply monitors Jack's brain activity, but it does not have any causal influence on what actually happens. Sam acts exactly as he would have acted, had no device been implanted in his brain.

Evidently, Sam is morally responsible for what he has done. Indeed, Sam is blameworthy for deciding to kill the mayor and for killing the mayor. But whereas Sam is morally responsible for his action, he could not have done otherwise. Sam could not have done otherwise because of the existence of a 'counterfactual inter-

[5] We introduce this example in J. M. Fischer and M. Ravizza, 'Responsibility and Inevitability'.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

vener' (Jack) who would have caused him (in a certain manner) to behave as he actually did, had Sam been inclined to do otherwise. We might say that Sam acts freely and is morally responsible for what he does because no 'responsibility-undermining factor' operates in the actual sequence leading to his action. Rather, such a factor — Jack's use of the electronic device to stimulate Sam's brain — operates in the *alternative* sequence. In cases in which a responsibility-undermining factor operates in the alternative sequence but not in the actual sequence, an agent can be held morally responsible for an action, although he could not have done otherwise. The case of Sam and Jack is such a case; let us call it 'Assassin'. 'Assassin' is a 'Frankfurt-type' case.[6]

So far we have been focusing upon moral responsibility for actions. We now wish to turn to the issue of moral responsibility for *consequences* of what we do. Before proceeding, we should point out that the events and states of affairs which are consequences of what we do can be construed as either particulars or universals. For our purposes, the distinction between consequence-particulars and consequence-universals will be made in terms of criteria of individuation. We shall stipulate that a consequence-particular is individuated more finely than a consequence-universal. Specifically, the actual casual pathway to a consequence-particular is an essential feature of it, so that if a different causal pathway were to occur, then a different consequence-particular would occur. In contrast, we stipulate that the same consequence-universal can be brought about via different causal antecedents.

For example, in 'Assassin' one can distinguish between the consequence-particular, *the mayor's being shot*, and the consequence-universal, *that the mayor is shot*. Had there been some indication that Sam would not shoot the mayor, and had Jack's device played a causal role in producing the outcome, a different consequence-particular would have occurred. (A different consequence-particular would have been denoted by, 'the mayor's being shot'.) In contrast, even if Jack's device had played a causal role, the same consequence-universal, *that the mayor is shot*, would have occurred. In the case of 'Assassin', Sam shoots the mayor without the intervention of Jack's electronic device. In this case, the state of affairs, *that the mayor is shot*, obtains. But this same state of affairs could have been caused to obtain in various different ways; in particular, it would have obtained even if Jack's device had caused Sam to shoot the mayor. Now the question arises as to whether an agent can be morally responsible for a consequence-universal, where the agent could not have prevented the consequence-universal from obtaining. Are there cases in which an agent is morally responsible for the occurrence of a consequence which is inevitable (for him)? (In what follows, we shall be primarily concerned with consequence-universals.)

It appears as if there are such cases. Take, for example, 'Assassin'. It is plausible to say that Sam is morally responsible not only for shooting the mayor, but also for the consequence-universal, *that the mayor is shot*. And note that Sam cannot prevent the mayor from being shot in one way or another. Presumably, also, Sam

[6] H. G. Frankfurt, 'Alternate Possibilities and Moral Responsibility', *The Journal of Philosophy* 66 (1969) pp. 828-839; and 'Freedom of the Will and the Concept of a Person', *The Journal of Philosophy* 68 (1971) pp. 5-20.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

can he held morally responsible for the state of affairs, *that the mayor is killed.* And note again that Sam cannot prevent the mayor from being killed in one way or another.

Heinaman articulates an account of moral responsibility for consequence-universals by reference to which one could say that Sam is indeed morally responsible for these inevitable states of affairs. (Heinaman does not discuss the specific case, 'Assassin', but his account is meant to apply to such a case.) On Heinaman's approach, an agent is morally responsible for a state of affairs insofar as he knowingly and intentionally does something the doing of which at least in part explains why the state of affairs obtains.[7] Heinaman points out that one can do something the doing of which is a sufficient condition for the obtaining of a certain state of affairs without its being the case that one's doing the thing in question even in part explains why the state of affairs obtains. So, for example, one can break a glass, which is presumably a sufficient condition for its being the case that the glass is brittle. But one's breaking the glass need not in any way explain why the glass is brittle. Similarly, Sam's shooting the mayor is a sufficient condition for the mayor's being mortal, but this fact does not entail that Sam's shooting the mayor even in part explains why the mayor is mortal. Thus, on Heinaman's account, one need not say that in 'Assassin' Sam is morally responsible for the state of affairs, *that the mayor is mortal.* But it *is* plausible on this account to say that Sam is morally responsible for the states of affairs, *that the mayor is shot,* and *that the mayor is killed*: Sam's action helps to explain why these states of affairs obtain.

Having provided an account of moral responsibility for consequence-universals which appears to have plausible results in such cases as 'Assassin', Heinaman turns to a defence of this account against a strategy of response he attributes to Peter van Inwagen. van Inwagen wishes to insist in general on the criterion of the existence of genuinely open alternative possibilities for moral responsibility. In particular, van Inwagen holds that in order for an agent to be morally responsible for a state of affairs (universal), the agent must be able to prevent the state of affairs from obtaining. Thus, van Inwagen must argue that there is something wrong with Heinaman's approach. Specifically, van Inwagen might argue as follows. In 'Assassin', Sam is *not* morally responsible for the state of affairs

> (K) The mayor is shot by Sam on his own or the mayor is shot by Sam after having been prompted by Jack's mechanism.

But if Sam is not morally responsible for (K), then under the circumstances of 'Assassin', Sam is not responsible for the state of affairs, *that the mayor is shot.*

But despite the surface plausibility of van Inwagen's argument, Heinaman claims (first) that an equally plausible argument can be constructed which shows that in 'Assassin' Sam *is* morally responsible for the state of affairs, *that the mayor is shot.* Pursuing this approach, Heinaman says that Sam is morally responsible (in

[7] *Ibid.*, p. 271.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

'Assassin') for the state of affairs, *that the mayor is shot by Sam on his own.* Now, given this point and

> (I) If S is responsible for a state of affairs A, then if 'F' is any false statement (other than 'not-A'), then S is responsible for the state of affairs that A or F,

it follows that Sam is responsible for the state of affairs, *that Sam shoots the mayor on his own or that Sam shoots the mayor after having been prompted by Jack's mechanism.* But if so, then given the situation in 'Assassin', Sam is morally responsible for the state of affairs, *that the mayor is shot.* We thus have an argument which is at least as plausible as van Inwagen's for the *opposite* conclusion—that Sam can indeed be morally responsible for a state of affairs which he cannot prevent from obtaining. Insofar as van Inwagen's argument is (at best) no more plausible than the counter-argument, Heinaman can claim that van Inwagen's argument does not refute Heinaman's account.

Further, Heinaman points out that the justification presented by van Inwagen for his claim that Sam is not morally responsible for

> (K) The mayor is shot by Sam on his own or the mayor is shot by Sam after having been prompted by Jack's mechanism.

is that (K) is not relevantly different from such states of affairs as

> (K*) The mayor is shot by Sam or 2 + 2 = 4, and
> (K**) The mayor is shot by Sam or grass is green.

But for the latter two states of affairs Sam is manifestly not morally responsible. Hence, it is alleged by van Inwagen that Sam is not morally responsible for (K). Against this Heinaman claims that there *is* a relevant difference between (K) and the latter two states of affairs — (K*) and (K**). Heinaman points out that whereas in (K) the second disjunct is false, in (K*) and (K**) the second disjuncts are true; thus, in virtue of principle (I), the first state of affairs can be distinguished from the latter two. Hence, Heinaman employs principle (I) to block van Inwagen's argument that one cannot be morally responsible for a state of affairs such as (K).

Further, Heinaman suggests that a plausible explanation of the intuitive judgment that Sam is not morally responsible for (K*) and (K**) employs

> (II) If S is not responsible for A, then for any proposition P, S is not responsible for the state of affairs that A or P.

Whereas one can use principle (II) to show why Sam is not morally responsible for (K*) and (K**), one cannot employ (II) to show that Sam is not morally responsible for (K). Thus, according to Heinaman, even if (I) were false, (II)'s truth implies that there is a relevant difference between such states of affairs as (K) and such states of affairs as (K*) and (K**).

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

II. Critique of Heinaman

We begin with a criticism of Heinaman's account of moral responsibility for consequence-universals: an agent is morally responsible for a state of affairs insofar as the agent knowingly and intentionally does something which at least in part explains why the state of affairs obtains.[8] Consider 'Missile'.[9] In 'Missile' Joan knows that Elizabeth has already launched a missile toward Washington D.C. But Joan has a weapon which she could use to deflect the missile in such a way that it would hit a less populous area of Washington D.C. Unfortunately, Joan is located very close to the city, and because of this fact, the bomb's trajectory, and the nature of her weapon, she knows that, whereas she can deflect the bomb onto a different part of the city, she cannot prevent the bomb from hitting the city at all.

Imagine that Joan does employ her weapon to deflect the bomb. When she does this, she may well be morally responsible for the fact that one section of Washington (rather than another) is bombed. But is Joan morally responsible for the fact that Washington D.C. is bombed? It seems that she is *not* (given that Joan had nothing to do with Elizabeth's launching the missile).

Heinaman's criterion now faces the following problem. Heinaman must say that in 'Missile' Joan is morally responsible for the state of affairs, *that Washington D.C. is bombed (somewhere or another)*. That is to say, *if* it is plausible to suppose that (in 'Assassin') Sam's act of pulling the trigger at least in part explains why the state of affairs, *that the mayor is killed (in some way or another)*, obtains, then it is equally plausible to suppose that (in 'Missile') Joan's act of deflecting the bomb at least in part explains why the state of affairs, *that Washington D. C. is bombed (somewhere or another)*, obtains. But it is an implausible result that Joan must be deemed responsible for this consequence-universal. Whereas it is reasonable to think that Joan is morally responsible for the fact that Washington D. C. is bombed *in one area rather than another*, it is implausible to say that she is morally responsible for the fact that Washington D. C. is bombed somewhere or another.

The criticism presented above provides the materials for a further criticism: it shows why Heinaman's Principle (I) is false. Recall

> (I) If S is responsible for a state of affairs A, then if 'F' is any false statement (other than 'not-A'), then S is responsible for the state of affairs that A or F.

Consider again 'Missile'. In this case Joan is morally responsible for the fact that the missile (already launched by Elizabeth) hits a particular part of Washington D.C. Now form the disjunction of the true statement about where the missile hits and all the false statements about where the missile hit the city of Washington. Joan is *not*

[8] We believe that an agent can knowingly and intentionally do something without *freely* doing it; under such circumstances, knowingly and intentionally doing something which at least in part explains why a certain state of affairs obtains is not sufficient for moral responsibility for that state of affairs. We shall not dwell on this objection here.

[9] We introduce this example in Fischer and Ravizza, 'Responsibility and Inevitability'. A similar example was originally suggested to us by David Widerker.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

morally responsible for this state of affairs. After all, this disjunction is equivalent to the fact, *Washington D.C. is bombed in some place or other*, and Joan is *not* morally responsible for this fact. Thus, Heinaman's (I) must be rejected.

Thus far we have presented an example which indicates that Heinaman's account of moral responsibility for consequence-universals is inadequate. This example also shows that Heinaman's (I) must be rejected. In the example there is some tendency to say that an agent is not morally responsible for the pertinent state of affairs precisely because he cannot prevent it from obtaining. Of course, this is exactly what van Inwagen (*pace* Heinaman) would insist. Further, principle (I) was invoked by Heinaman to undermine van Inwagen's argument that one cannot be responsible for a state of affairs which one cannot prevent from obtaining; insofar as (I) has been shown to be false, there is further pressure toward van Inwagen's position.

Heinaman suggests that even if (I) were false,

> (II) If S is not responsible for A (an existing state of affairs), then for any proposition P, S is not responsible for the state of affairs that A or P

is secure and can be employed to undermine van Inwagen's argument for the conclusion that an agent can be morally responsible for a state of affairs only if he can prevent that state of affairs from obtaining. Specifically, as argued above, one can adduce (II) to distinguish between such states of affairs as (K) and such states of affairs as (K*) and (K**). (Of course, if one cannot distinguish these kinds of states of affairs, there is more pressure to adopt van Inwagen's view.)

What lies behind the apparent intuitive plausibility of principle (II)? One might appeal to the following sort of principle:

> (J) If one is not responsible for one state of affairs, and this state of affairs is all that is required for a second state of affairs, then one is not responsible for the second state of affairs.

Presumably, if (J) were true, then one could defend (II). But (J) is not true. We can see this by again considering Frankfurt-type cases such as 'Assassin'. In 'Assassin' Sam is not morally responsible for the existence of Jack's mechanism which is, under the circumstances, sufficient for the state of affairs that the mayor is shot. But it is intuitively plausible that Sam is nevertheless morally responsible for the state of affairs that the mayor is shot.

Frankfurt-type cases such as 'Assassin' involve *preemptive* overdetermination. There are other cases which involve *simultaneous* overdetermination which again point to the falsity of (J). Consider a case which usually falls under the rubric of 'collective responsibility'.[10] In this example, 'Joint Assassins', Charlie wants to assassinate the president of the United States. As the president steps to the podium at T1, Charlie aims at her head and fires a shot. At precisely the same moment, another assassin, Betty, also fires a shot at the president's head from another part of

[10] The example is borrowed from G. Mellema, 'On Being Fully Responsible', *American Philosophical Quarterly* 21 (1984) pp. 189-194, esp. p. 190.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

the crowd. Both bullets find their mark simultaneously, and the president is killed at T2. Neither assassin knew the other assassin was present, and either bullet alone would have been sufficient to kill the president. In this case, it seems natural to suppose that both persons are at least partly responsible for the consequence-universal, *that the president is killed at T2*. Each agent may in fact be fully (although not solely) responsible for this consequence, but at the very least each assassin is partly responsible for it.

Note that it would be implausible to argue that Charlie is not even partly responsible for the president's being killed at T2 in virtue of the presence (and intentions) of Betty. By symmetry one could argue that Betty is not even partly responsible for the president's being killed at T2 in virtue of the presence (and intentions) of Charlie. Thus, it would follow that *neither* of them (and thus, presumably, no one) is even partly responsible for the president's being killed at T2 — a manifestly implausible conclusion. Thus, in 'Joint Assassins', Charlie is not responsible for a state of affairs (Betty's presence and intentions) which is sufficient for a second state of affairs (the president's being killed at T2), and nevertheless Charlie is at least partly responsible for the second state of affairs. If Charlie is fully (although of course not solely) responsible for this latter state of affairs, then (J) must be false. And if (J) is false, it is unclear how one might justify principle (II). Indeed, we believe that the above sorts of cases show that (II) is false.

It might be worthwhile to point out that in cases such as 'Assassin' and 'Joint Assassins', another principle seems applicable (and seems at first at least as plausible as [J]):

> (J*) If an agent is morally responsible for one state of affairs and this is all that is required for a second state of affairs, the agent is morally responsible for the second state of affairs.

More will be said about (J*) below, but it suffices here to point out that it is intuitively at least as plausible as (J) and equally applicable to the relevant cases.

So the dialectical situation here is as follows. One might be attracted to a principle such as (J) to explain and buttress (II). But (J) is shown to be false by both cases of preemptive and simultaneous overdetermination. Further, (J*) is — on the surface — at least as attractive as (J). And (J*) can be employed to show the *falsity* of (II). It suffices here to notice that one cannot simply invoke (J) to support (II), given the existence and symmetric attraction of (J*).

III. An Alternative Approach

We have argued that Heinaman's account of moral responsibility for consequence-universals is inadequate. The kinds of examples we invoked in this criticism also imply that Heinaman's Principle (I) is false. Further, we have argued against Principle (II). Thus, Heinaman has provided no good reason to reject van Inwagen's claim that moral responsibility for a consequence-universal requires the ability to prevent that state of affairs from obtaining. Indeed, van Inwagen's view is

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

supported by examples such as 'Missile'.

But we have claimed that examples such as 'Assassin' show that moral responsibility for a consequence-universal need *not* require the ability to prevent that state of affairs from obtaining. Thus, whereas Heinaman's approach leads to implausible results in cases such as 'Missile', van Inwagen's approach leads to implausible results in cases such as 'Assassin'. *Neither* approach is satisfactory. What is needed is a new approach which says the correct thing about *both* sorts of case. We shall now undertake to sketch (in an admittedly brief fashion) such an approach, after which we shall apply it to the examples discussed above.

III.1. *Responsibility for actions.* We shall here present just the barest sketch of a theory of moral responsibility for actions. This theory has been elaborated elsewhere in greater detail, and it must be further refined and developed in future work.[11] But it is useful here to see how our theory of moral responsibility for consequences builds upon and fits together with the theory of responsibility for actions.

The basic idea is that (given that the relevant epistemic conditions are met[12]) an agent is morally responsible for performing an action insofar as i) it is not the case that the agent does what he does solely because he could not have done otherwise, and ii) the bodily movement identical to the action issues from a weakly reasons-responsive mechanism. We shall here give a very brief sketch of weak reasons-responsiveness. To say whether an action issues from a weakly reasons-responsive mechanism, we first need to identify the kind of mechanism that actually issues in action. It is important to see that, in some cases, intuitively different kinds of mechanisms operate in the actual sequence and the alternative sequence. So, for instance, in 'Assassin' the ordinary process of practical reasoning issues in Sam's act of squeezing the trigger, but a different type of mechanism (involving direct electronic stimulation of Sam's brain) would have operated, had Sam shown any sign of wavering. We cannot here develop an explicit account of mechanism-individuation. It suffices, for our purposes, to note that Sam's actual-sequence mechanism is intuitively of a different sort from the alternative-sequence mechanism. Whereas it is difficult to produce an explicit criterion of mechanism-individuation, we believe that it is natural to say that in a case such as 'Assassin' different sorts of mechanisms issue in the actions in the actual and alternative sequences.

In order to determine whether an actual-sequence mechanism of a certain type is weakly reasons-responsive, one asks whether there exists some possible scenario (with the same natural laws as the actual world) in which that type of mechanism operates, the agent has reason to do otherwise, and the agent does otherwise (for that reason). That is, we hold fixed the actual type of mechanism, and we ask whether the agent would respond to some possible incentive to do otherwise. If so, then the

[11] Fischer, 'Responsiveness and Moral Responsibility'. See also, M. Ravizza, "Is Responsiveness Sufficient For Responsibility?', (typescript, University of California, Riverside), and Fischer and Ravizza, 'Responsibility and Inevitability'.

[12] Following Aristotle, one might distinguish between an epistemic and a freedom-relevant component of a theory of responsibility. Aristotle held that one acts voluntarily insofar as one is not in a relevant sense ignorant of what one is doing and one is not compelled to do what one does. We focus here primarily on the freedom-relevant component. For a useful discussion of issues pertinent to the epistemic component, see J. Feinberg, *Harm to Self* (Oxford: Oxford University Press, 1986) pp. 269-315.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

actually operative mechanism is weakly reasons-responsive. In contrast, strong reasons-responsiveness obtains when a certain kind K of mechanism actually issues in an action, and if there were sufficient reason to do otherwise and K were to operate, the agent would recognize the sufficient reason to do otherwise and thus choose to do otherwise and do otherwise.

III.2. *Responsibility for consequences.* The account of moral responsibility for consequences is in certain respects parallel to (and also an extension of) the account of moral responsibility for actions.[13] The leading idea is that (given that certain epistemic constraints are satisfied) an agent is morally responsible for a consequence insofar as i) it is not the case that the agent brings about the consequence solely because he believes that he cannot prevent it from occurring (in some way or another), and ii) the consequence emanates from a responsive *sequence*. It is necessary to distinguish *two components* of the sequence leading to a consequence. The first component is the mechanism leading to action (bodily movement), and the second component is the process leading from the action to the event in the external world. We shall say that, in order for the sequence leading to a consequence to be responsive, both the mechanism leading to the action must be weakly reasons-responsive and the process leading from the action to the consequence must be sensitive to action'.

Before proceeding to the account, it is important to note that the counterfactual intervener in a Frankfurt-type case need not be another agent (whose action in the alternative sequence would bring about the consequence in question). As Frankfurt points out, the role of counterfactual intervener may be played 'by natural forces involving no will or design at all'.[14] Given this, it seems that in evaluating the sensitivity of a process one wants to hold fixed not only the actions of other agents in the actual sequence, but also any natural events which play no role in the actual sequence but which would, in the alternative sequence, *trigger* causal chains leading to the consequence in question. For convenience we can group *both* other actions that would trigger causal chains leading to the consequence *and* natural events that would do so under the heading, 'triggering events'. Let us think of a triggering event (relative to some consequence C) as an event which is such that if it were to occur, it would *initiate* a causal sequence leading to C.

Suppose that in the actual world an agent S performs some action A via a type of mechanism M , and S's A-ing causes some consequence C via a type of process P. We shall say that the sequence leading to the consequence C is responsive if and only if there exists some action A^* (other than A) such that: (i) there exists some possible scenario in which an M-type mechanism operates, the agent has reason to do A^*, and the agent does A^*; and (ii) if S were to do A^*, all triggering events which do not actually occur were not to occur, and a P-type process were to occur, then C would not occur.[15]

Before showing how this principle can be applied to explain our intuitive judgments about the cases discussed above, let us take a moment to discuss a few points

[13] See Fischer and Ravizza, 'Responsibility and Inevitability'.

[14] Frankfurt, 'Alternate Possibilities and Moral Responsibility', fn. 4.

[15] It should be noted that the analysis of a responsive sequence given above is for cases of individual responsibility; it is not intended to address cases of simultaneous overdetermination in

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

which should help both to clarify and to illustrate our principle. First, in formulating the definition of a responsive sequence, we make use of the intuitive notion of a 'type of process' leading from the action to the event in the external world. This is parallel to the notion of a kind of mechanism issuing in action. As above, we concede both that process-individuation might be problematic and that we do not have an explicit theory of process-individuation. But, as above, we believe that there is a relatively clear intuitive distinction between different types of processes.

We do not deny that there will be difficult questions about process-individuation. Nevertheless, all that is required for our purposes is that there be agreement about some fairly clear cases. If we are unsure about an agent's moral responsibility for a consequence in precisely those cases in which we are unsure about process-individuation, then at least the vagueness in our theory will match the vagueness of the phenomena it purports to analyze.

Second, in ascertaining the responsiveness of a particular sequence involving a mechanism issuing in action, an action, and a process leading from the action to a consequence, we 'hold fixed' the type of mechanism and the type of process. If it is the case that a different mechanism or process would have taken place if things had been different (i.e., if the case is a 'Frankfurt-type' case), this is irrelevant to the responsiveness of the *actual* sequence. Further, imagine that we are testing the sensitivity of a particular process leading from an action to a consequence. Suppose that the agent actually performs a certain action thus causing some consequence, and that no one else actually performs that type of action. Under these conditions, we 'hold fixed' others' behaviour when we test for the sensitivity of the process leading from action to consequence. The point is that, when we are interested in the sensitivity of the process to action, we are interested in whether there would have been a different outcome, if the agent had not performed a certain sort of action *and all non-occurring triggering events were not to occur.*

The theory claims that the sequence leading to a consequence includes more than just the mechanism issuing in action. Thus, it is not surprising that both components — i.e., the mechanism leading to the action and the process leading from the action to the event — are relevant to responsibility for a consequence, where only the first component is relevant to responsibility for an action. Thus, the theory of responsibility for a consequence involves two stages. It will be seen below that this two-stage approach helps us appropriately to distinguish different cases of responsibility for consequences which are inevitable. Further, it is important to note that when considering responsibility for consequences the second component should not be considered in isolation from the first. Our definition of the responsiveness of the sequence leading to a consequence requires a certain sort of *linkage* of the two components of the sequence.

[15] *continued. . .*

which several agents may be jointly responsible for the consequence produced. We believe that a similar analysis may work in these case, but in evaluating the responsibility of any one agent it might be necessary to 'bracket' the other triggering events that simultaneously produce the consequence, in order to ascertain if the agent's action was part of a responsive sequence that was sufficient to produce the consequence. The issues involved in such cases of joint responsibility are complex and we cannot fully pursue them in the present essay.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

Third, the notion of a 'triggering event' is — like the notions of 'mechanism' and 'process' — fuzzy around the edges. But, again, we believe that it is tolerably clear for the present purposes. Note that a triggering event is an event which would 'initiate' a causal chain leading to a certain consequence. Although the concept of 'initiation' is difficult to articulate crisply, we rely on the fact that there are some fairly uncontroversial instances of the concept. So, for example, if a lightning bolt hits a house and there is a resulting fire, the event of the lightning's hitting the house could be said to initiate the sequence leading to the destruction of the house. And this is so even if there were certain atmospheric events which antedated the lightning bolt and which led to it. Of course, the notion of 'initiation' is highly context-dependent, and the truth of claims about purported initiations will depend on the purposes and goals of the individuals making (and considering) the claims. But we believe that the notion of initiation issues in tolerably clear intuitive judgments about the cases relevant to our purposes.

III.3. *Applications.* With the principle of moral responsibility for consequences in hand, we can explain the intuitive judgments about cases described above. In 'Assassin' Sam is morally responsible for the consequence-universal, *that the mayor is killed.* In this case, the actual-sequence mechanism (ordinary practical deliberation) is weakly reasons-responsive, and the process leading from action to consequence (ordinary physical laws, no 'abnormal circumstances') is sensitive to action: had Sam not squeezed the trigger (either as a result of his own deliberation or because of Jack's intervention) and others' relevant behaviour were held fixed, the mayor would not have been killed. Thus, the two components necessary for actual causal control are present, and Sam can be held morally responsible for the fact that the mayor is killed, although he could not have prevented it.

But in 'Missile' Joan is not morally responsible for the consequence-universal, *that Washington D.C. is bombed.* This is because the sequence including Joan's action and the process leading from her action to the event of Washington's being bombed is not responsive. Of course, the first component *is* weakly reasons-responsive (and thus Joan can be held morally responsible for her *action* of deflecting the bomb). But the sequence is not responsive, because the second component — the process leading from action to event in the world — is *not* sensitive to action. That is, the world is such that, no matter how Joan acts, the bomb will hit Washington D.C., given that the appropriate actual-sequence facts are held fixed.

Our principle of moral responsibility for consequences, then, explains the intuitive judgments about various examples presented above. Further, the theory explains why there is an important difference between such cases as 'Assassin' and such cases as 'Missile'. The agent could have prevented the relevant consequence-universal in *none* of these cases. But, whereas in 'Assassin' the relevant consequence-universal issues from a responsive sequence, in 'Missile' the consequence-universal does *not* issue from a responsive sequence.

Our theory will allow us to analyze such states of affairs as

> (K) The mayor is shot by Sam on his own or the mayor is shot by Sam after having been prompted by Jack's mechanism,

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

(K*) The mayor is shot by Sam or 2 + 2 = 4, and
(K**) The mayor is shot by Sam or grass is green

in a perspicuous way. We agree with Heinaman (*pace* van Inwagen) that the three states of affairs should not be assimilated in regard to the pertinent responsibility ascriptions. But recall that Heinaman invokes principle (I) as his instrument of differentiation, claiming that (I) distinguishes the three states of affairs in virtue of the asymmetric truth values of the second disjuncts in the trio of propositions. But we have argued that Heinaman's (I) is false.

On our approach, (K), (K*), and (K**) are indeed different. All three states of affairs are impossible for Sam to prevent. All three are inevitable: all three will obtain, no matter how Sam moves his body, among all they ways Sam can move his body. But, as we proceed to argue, (K) — as opposed to the other states of affairs — is responsive relative to the relevant sequence. Thus, the resources of our theory generate a mechanism of discrimination of the three states of affairs and hence insulate the theory from the van Inwagen-type objection considered by Heinaman *without* appeal to (I).

Before elaborating, we pause to point to an incompleteness in our adumbration of an account of moral responsibility for consequence-universals. We pointed out that the sequence issuing in a consequence has *two* stages: the stage leading to the action (bodily movement) and the stage leading from the action to some event in the 'external' world (apart from the agent). This picture applies most naturally to 'simple' consequence-universals. But there can also be various sorts of complex consequence-universals (states of affairs) to which the theory should apply. Thus, in some consequences there are not only two stages, but two (or more) components of the second stage.

The trio of propositions, (K), (K*), and (K**) correspond to just these kinds of complex states of affairs. In each member of the trio of states of affairs, there is some event which is the terminus of the second stage of the relevant sort of consequence-producing sequence *disjoined with* another event or state of affairs. The added disjunct renders the disjunctive state of affairs complex. It is interesting that the second disjunct in all of these states of affairs has an actual or possible source which is *different from* the source of the event which is the terminus of the second stage of the consequence-producing sequence; we suppose that there could be complex states of affairs whose various different 'second-stage' elements all issue (perhaps even in interestingly different *ways*) from the *same* first-stage element. These complex states of affairs would have a 'branching' structure; we shall put these consequence-universals aside here, postponing our discussion to another occasion.

Aside from such states of affairs with a branching structure, it is evident that there are various importantly different sorts of states of affairs whose second-stage elements come from *different* first-stage elements. (K*) and (K**) contain second-stage elements that come from different first stage-elements, one of which does not involve human agency. Other such states of affairs may contain second-stage elements which come from human agency. In certain cases of this sort, the relevant agent is in control of the other agent (or agents); in other such cases, the relevant

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

agent is *not* in control of the other agent (or agents). Again, we shall not here attempt to address the full range of relevant possibilities.

Although discussion of the appropriate application of our approach to the full range of different kinds of complex consequence-universals is beyond the scope of this paper, we shall show how we would apply the account to such states of affairs as (K), (K*), and (K**); in so doing, we shall *extend* the account (as presented above) to apply to complex states of affairs of a certain sort. The main point for our purposes here is that it is natural and intuitive to apply our account in such a way that it implies that (K) is associated with a responsive sequence, whereas neither (K*) nor (K**) is. (K) is interestingly different from (K*) and (K**) to the extent that (K) has only one true disjunct whereas (K*) and (K**) have two true disjuncts; thus, whereas it is intuitively relatively easy to say what 'makes true' or 'brings about' (K), it is a more delicate matter to say what 'makes true' or 'brings about' (K*) and (K**). (The latter states of affairs are the 'consequence-analogues' of cases of simultaneous overdetermination of *actions*.) So, on our approach, it is relatively easy to say that (K) *issues from* a responsive sequence: but, given that it is unclear what 'makes true' or 'brings about' (K*) and (K**), it is more difficult to ascertain their status with regard to responsiveness.

But there is a natural and intuitive way of resolving this problem. In all three members of the trio of states of affairs, there is some concrete action by Sam, and then the larger state of affairs is constructed out of this element and additional components. In all three cases, we first have Sam's act and then we ask whether Sam is morally responsible for the larger state of affairs *in virtue* of his concrete action (and its result in the world). Thus, it is natural and appealing to take the sequence issuing in the 'first disjunct' — the sequence which involves Sam's action — to be the *relevant* sequence with respect to which responsiveness of the larger state of affairs is assessed. Thus, there are two ways in which a state of affairs may be responsive or associated with a responsive sequence: it may *issue from* a responsive sequence, or it may be responsive *relative to the relevant sequence*.

Now we can put our point more sharply. Whereas (K) is associated with a responsive sequence, neither (K*) nor (K**) is associated with a responsive sequence. Holding fixed the pertinent properties of the sequence involved in the first disjunct and varying the reasons, there are scenarios in which (K) is false; but holding fixed those properties and varying the reasons, there are *no* scenarios in which (K*) or (K**) are false. Thus, a natural application of our approach to such complex consequence-universals as (K), (K*), and (K**) implies the appropriate results: (K) is associated with a responsive sequence, whereas neither (K*) nor (K**) is.[16]

[16] It might be thought that employing our theory in the way just sketched will have the counterintuitive result that neither of the agents in 'Joint Assassins' is even partly responsible for the upshot, *the president's being shot*. But this is not in fact so, because in a case of simultaneous overdetermination — in which two causal sequences actually operate and issue in the same concrete event (in virtue of which the relevant state of affairs obtains) — one must 'bracket' the other sequence when ascertaining the responsiveness of each individual sequence. (See fn. 15.) In the absence of such overdetermination, bracketing is inappropriate.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

Finally, let us return to the critique of principle (II) and the attendant discussion of the ancillary principles (J) and (J*). We argued that certain cases (such as 'Assassin') show (II) to be false. But it might appear that our method involved starting with the claim that an agent is responsible for some state of affairs (say, *shooting the mayor on one's own*) and then claiming that the agent must then be responsible for the disjunction of this state of affairs with any other state of affairs (say, *shooting the mayor having been prompted to do so by Jack's mechanism*). But if this strategy lies behind our rejection of (II), can't it be employed to block our rejection of (I)? That is, if Joan is morally responsible for the missile's hitting Washington D.C. in one particular area, then the strategy apparently could be employed to show that she is responsible for the missile's hitting Washington D.C. in either that area or any of the other areas. And if this is so, it would seem that we would have derived the repugnant conclusion that in 'Missile' Joan is morally responsible for the state of affairs, *that Washington D.C. is bombed somewhere or other*. Apparently, our strategy for rejecting (II) undermines our strategy for rejecting (I).

But the appearance is misleading, and again our sketch of a theory of moral responsibility for consequence-universals provides the resources which can be used to explain why this is so. In 'Assassin' Sam is morally responsible for the state of affairs, *that the mayor is shot*; this is because in 'Assassin' this state of affairs issues from a responsive sequence. That is to say, the state of affairs, *Sam shoots the mayor on his own or Sam shoots the mayor after having been prompted by Jack's mechanism*, issues from a certain sort of responsive sequence. Thus, when an agent is responsible for a state of affairs which is disjoined with a second state of affairs to produce a disjunctive state of affairs, the agent can be held morally responsible for the disjunctive state of affairs insofar as the disjunctive state of affairs is associated with a responsive sequence (of a certain sort). Note that this is the case in 'Assassin', but *not* in 'Missile'. That is, even though the state of affairs, *Washington D.C. is bombed in this particular place*, issues from a responsive sequence, the disjunctive state of affairs equivalent to the state of affairs that Washington D.C. is bombed somewhere or other is *not* associated with a responsive sequence. Thus, the strategy we employed to reject (II) — when properly understood — need not undermine the strategy we employed to reject (I).

The above remarks suggest that we ought to reconsider the alleged justification of the rejection of (II),

> (J*) If an agent is morally responsible for one state of affairs, and this is all that is required for a second state of affairs, the agent is morally responsible for the second state of affairs.

This principle should be revised to

> (J**) If an agent is morally responsible for one state of affairs (in virtue of its issuing from a responsive sequence S), and this is all that is required for a second state of affairs (which itself is responsive relative to S), the agent is morally responsible for the second state of affairs.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

If (J**) supports one's rejection of (II), then one need not be pushed to accept (I). Insofar as responsibility is associated with *responsiveness*, we have an understanding of these superficially puzzling phenomena, and we can analyze the pertinent examples in a penetrating fashion.

IV. Conclusion

Some cases — such as 'Missile' — lead one to think that in order to be morally responsible for a consequence-universal, one must be able to prevent it from obtaining. Other cases — such as 'Assassin' — make it tempting to think that one can be morally responsible for a consequence-universal which one cannot prevent from obtaining. We have sought to sketch an account which says the intuitively correct thing about both sorts of case. On this account, one is not responsible for the pertinent consequence-universal in 'Missile', but *not* because such responsibility requires the ability to prevent it from obtaining. Thus, it is open to us to say the agent *is* responsible for the relevant state of affairs in 'Assassin'. Our approach develops a certain association of moral responsibility with *responsiveness*.

Notice that principles

> (J) If one is not responsible for one state of affairs, and this state of affairs is all that is required for a second state of affairs, then one is not responsible for the second state of affairs, and
>
> (J*) If an agent is morally responsible for one state of affairs and this is all that is required for a second state of affairs, the agent is morally responsible for the second state of affairs,

are 'warring principles'. That is, they lead to opposite conclusions in certain cases, despite their surface plausibility. Take 'Assassin'. If one applies (J), one can generate the conclusion that Sam is *not* morally responsible for the consequence-universal, *that the mayor is shot*. That is, if one starts with (J) and focuses on the fact that Sam is not morally responsible for the existence of Jack's mechanism, one can be lead to conclude that Sam is not morally responsible for the fact that the mayor is shot. And if one applies (J*), one can generate the conclusion that Sam *is* morally responsible for the consequence-universal, *that the mayor is shot*. That is, if one starts with (J*) and focuses on the fact that Sam is morally responsible for his concrete action, one can conclude that Sam is morally responsible for the fact that the mayor is shot.

(J*) appears to provide some intuitive support for

> (I) If S is responsible for a state of affairs A, then if 'F' is any false statement (other than 'not-A'), then S is responsible for the state of affairs that A or F.

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011

And (J) appears to provide intuitive support for

> (II) If S is not responsible for A, then for any proposition P, S is not responsible for the state of affairs that A or P.

Thus, it is not at all surprising that (I) and (II) are also warring principles. In 'Assassin', accepting (I) and focusing on Sam's responsibility for his act of shooting the mayor will lead one to conclude that Sam *is* responsible for the fact that the mayor is shot. But accepting (II) and focusing on Sam's lack of responsibility for Jack's mechanism will lead one to precisely the opposite conclusion: that Sam is *not* morally responsible for the fact that the mayor is shot.[17]

Happily, we have shown that *both* (J) and (J*) are false — as are Heinaman's (I) and (II). Our sketch of an account of responsibility for consequence-universals provides the materials for a revised principle,

> (J**) If an agent is morally responsible for one state of affairs (in virtue of its issuing from a responsive sequence S), and this is all that is required for a second state of affairs (which itself is responsive relative to S), the agent is morally responsible for the second state of affairs.

If (J**) is indeed true, this would support our contention that in 'Assassin' Sam is morally responsible for the consequence-universal, *that the mayor is shot,* without undermining our contention that in 'Missile' Elizabeth is *not* morally responsible for the consequence-universal, *that Washington D.C. is bombed.* Further, as far as we can see, (J**) is irenic: it is true and 'at war' with no other true principle.

Finally, we again emphasize that our approach to moral responsibility for consequence-universals needs to be developed, and defended at various points. But perhaps the discussion above will provide more impetus for this sort of project, insofar as the approach appears to have felicitous and illuminating consequences. Although the theory must be given more concrete expression, we believe that this project is potentially fruitful at least in part because the theory helps us to achieve a more precise understanding of a puzzling set of phenomena.

University of California, Riverside

Received July 1991
Revised October 1992

[17] Given his acceptance of (I), this is further evidence that Heinaman cannot embrace (II).

Downloaded By: [University of Notre Dame] At: 20:43 24 February 2011