Saying Good-bye to the Direct Argument
the Right Way

Michael McKenna
Florida State University

In “The Incompatibility of Responsibility and Determinism,” Peter van Inwagen (1980) introduced the alluring Direct Argument. The Direct Argument tells us that moral responsibility is incompatible with determinism because, at a determined world, nonresponsibility for the facts of the past and the laws of nature results in nonresponsibility for what one does. The Direct Argument differs from other arguments for the same conclusion in that it does not attempt to show that determinism is incompatible with some feature perhaps necessary for moral responsibility, such as the ability to do otherwise. Rather, it fixes directly on moral responsibility and argues that, under the assumption of determinism, no one is morally responsible. How so? Nonresponsibility transfers across deterministic relations. If the facts of the past and the laws of nature entail every truth about what one does, as the thesis of determinism would have it, and no one is even in part morally responsible for the past and the laws, then no one is even in part morally responsible for what one does.

I would like to thank Eleonore Stump and David Widerker for illuminating conversations on this topic in years past. I presented a version of this essay to the Philosophy Department at Florida State University and profited from many insightful questions from the audience. A special note of thanks is owed to Joshua Gert, who offered several outstanding points. Finally, for their extensive and thoughtful written comments, I am indebted to Bernard Berofsky, Randolph Clarke, John Martin Fischer, Carl Ginet, Ishtiyaque Haji, Alfred Mele, Derk Pereboom, as well as the editors of and three referees for Philosophical Review.
The Direct Argument has drawn various critics, who have attempted to produce counterexamples to the inference principle that is at its core. In the following pages, I consider two notable efforts, one by John Martin Fischer and Mark Ravizza (1998), and another by David Widerker (2002). Each effort, I maintain, while compelling, is not sufficient to reject the Direct Argument. I then proceed to challenge the Direct Argument in a novel fashion. Van Inwagen, I argue, has not given us good reason to think that the inference principle in question has an adequate anchor in our inferential practices, especially in the light of the dialectical context with the compatibilist. Hence, it is not the compatibilist’s burden to produce counterexamples to the contested principle. As I see it, it is van Inwagen’s burden to produce relevant confirming instances of it. Assuming, as van Inwagen does, that the principle has an anchor in our patterns of reasoning implicates advocates of the Direct Argument in a highly contentious assumption, one that requires argumentation, where currently the Direct Argument provides none.

1. The Direct Argument

The Direct Argument is structurally similar to the better known Consequence Argument. It will be helpful to start by explaining the latter. The Consequence Argument seeks to establish the incompatibility of determinism and moral responsibility “indirectly” by establishing the incompatibility of determinism and something that is required for moral responsibility. Assuming that the pertinent sort of freedom required for moral responsibility involves the freedom to do otherwise, the Consequence Argument concludes that, if determinism is true, no agent is able to render false any true proposition regarding what he or she does at any moment in time. Hence, no agent is able to do otherwise. To get this conclusion, the Consequence Argument relies on some variation of a “powerlessness” modal operator regarding propositions that a person at a time is powerless to render false. Treating ‘N’ as the modal operator, \( p \) as a proposition, \( a \) as a person, and \( t \) as a time, one can read ‘\( Np_{a,t} \)’ as “\( p \) is true, and \( a \) at \( t \) is powerless to render \( p \) false.” For ease of

---

1. The Consequence Argument was first introduced by Carl Ginet (1966). Later David Wiggins (1973) presented a version, and then van Inwagen (1975) offered a formulation and further refinement. Because I will be shifting from van Inwagen’s formulation of the Consequence Argument to his parallel formulation of the Direct Argument, my presentation of the former will track his.
expression, in the remainder of this essay, I'll suppress the notional indexing to persons and times.

According to van Inwagen (1983, 94), one can derive true claims of powerlessness by relying upon an inference rule labeled:

\[ \square p \vdash \neg p. \]

Treating \( \square \) as expressing “broadly logical necessity,” and applying Rule \( \alpha \), one can say that for any person at any time, a truth such as, for example, a truth of mathematics, is such that every person is powerless to render it false.

The Consequence Argument also relies upon an inference principle that “transfers” powerlessness from one fact to another that is entailed by the original. To formalize the inference, van Inwagen (ibid.) introduced:

\[ \neg (p \supset q), \neg p \vdash \neg q. \]

Together with Rule \( \alpha \), this rule closes claims of powerlessness under entailment relations.

The Consequence Argument proceeds by stipulating that determinism is true. So let us define determinism as the thesis that the facts of the past and the laws of nature entail every truth about the future. More cautiously, determinism can be defined as:

\[ \square (P \& L \supset F). \]

Add two more assumptions in the form of principles expressing metaphysical truths, the Principle of the Fixity of the Past, \( FP \), and the Principle of the Fixity of the Laws, \( FL \). The former, \( FP \), asserts that for any person, the past is fixed relative to the present; he or she is powerless to render false any truth about the past. The latter, \( FL \), asserts that, for any person, the laws of nature are fixed; it’s not up to the person whether or not the laws obtain. He or she is powerless to render false any law of nature.

With the above resources, the Consequence Argument can now be employed to show that any action performed at a determined world is one that is unavoidable for the agent. We need an action that is a reasonable candidate for one that is freely willed, and for which the agent might be morally responsible (meriting either praise or blame). So suppose that world \( W \) is determined, and that at it, at noon Eastern
Standard Time, on Wednesday, March 9, 2005, the following proposition is true: George Bush begins to move his body in such a way as to start to do the dance, the hustle. Call this proposition s, which we can express in shorthand as *George Bush gets down*. Here, finally, is the argument fully formalized:

1. $\Box (P \land L \supset s)$  
   Assume determinism; apply it to s.
2. $\Box (P \supset (L \supset s))$  
   1, and Logic.
3. $N (P \supset (L \supset s))$  
   2, Rule $\alpha$.
4. $N P$  
   Assume FP.
5. $N (L \supset s)$  
   3, 4, Rule $\beta$.
6. $N L$  
   Assume FL.
7. $Ns$  
   5, 6, Rule $\beta$.

This argument tells us that, at the determined world W, George Bush is unable to render it false that he gets down. Hence, it is unavoidable for him that he gets down.

For the conclusion $Ns$ to be false—for it to be true that George Bush could have avoided getting down—one of three possibilities must be true. Either, contrary to premise 4, he is able to act in such a way that, if he were to so act, the facts of the past would be altered (Saunders 1968); or, contrary to premise 6, he must be able to act in such a way that, if he were to so act, the laws of nature that do obtain at W would not (Lewis 1981); or Rule $\beta$ is invalid, and therefore one cannot infer the unavoidability of what one does do from the unavoidability of the facts of the past, FP, the laws of nature, FL, and the truth of determinism (Slote 1982). Each of these options has been carefully explored by many thoughtful philosophers, and I do not wish to pursue here whether the Consequence Argument is unsound for any of these reasons. I mean only to call attention to the structure of the argument and reveal the different places one might press upon it.

Before turning to the Direct Argument, it will be helpful to make perspicuous why van Inwagen regards the Consequence Argument as an indirect one. Notice that, to get from the conclusion above to an incompatibility between moral responsibility and determinism, further premises are needed. I think it best to add two. First, persons act of their own free will only if they could have done otherwise. Second, persons are morally responsible for what they do only if they act of own free will. Call each of these, respectively, the *Alternative Possibilities*
Saying Good-bye to the Direct Argument the Right Way

Condition (AP), and the Free Will Condition (FW). The Consequence Argument does not directly establish that moral responsibility is incompatible with determinism. It argues that avoidability is incompatible with determinism, and then, assuming something like the AP Condition and the FW Condition set out here, the desired conclusion is achieved, taking two extra steps. The Direct Argument, by contrast, does not need these further steps.

Turning to the Direct Argument, it exploits a modal operator and inference principles similar to the ones used in the Consequence Argument. The Direct Argument relies on a “nonresponsibility” modal operator regarding propositions for which a person is not morally responsible. Treating ‘NR’ as the modal operator, \( p \) as a proposition, \( a \) as a person, and \( t \) as a time, one can read ‘NR\( p,a,t \)’ as “\( p \) is true, and \( a \) at \( t \) is not even partly morally responsible for the fact that \( p \).” Here again, for ease of expression, I’ll suppress the notational indexing to persons and times.

As for the inference principles, as van Inwagen set them out, they are treated as near cousins of Rule \( \alpha \) and Rule \( \beta \):

Rule A: \( \Box p \vdash NRp. \)
Rule B: NR (\( p \supset q \)), NR\( p \vdash NRq \).

According to Rule A, any broadly logically necessary truth \( p \) (such as that \( 2 + 2 = 4 \)) entails that no one is even partly morally responsible for it. According to Rule B, if no one is even partly morally responsible for some fact, \( p \), and no one is even partly morally responsible for the fact that \( p \) materially implies \( q \), then no one is even partly morally responsible for the fact that \( q \).

2. This formulation helps to make clear the proper results of Frankfurt’s argument against the Principle of Alternative Possibilities (PAP): one is morally responsible for what one has done only if one could have done otherwise (Frankfurt 1969). Notice that PAP can be directly inferred from the AP and the FW conditions. Failing to distinguish between FW and AP, some write as if the soundness of Frankfurt’s argument would warrant the denial of FW. I find this highly implausible since moral responsibility does require some sort of freedom or control condition. FW can simply be identified with that. Note that it is the assumption that AP is necessary for FW (or even identical with it) that would lead one to make this mistake. My suspicion is that philosophers who think about the relationship between AP and FW in this mistaken way do so because they simply take it as a definitional matter that free will is (or requires) alternative possibilities. But if Frankfurt’s argument can be used to show that moral responsibility does not require alternative possibilities, one could easily use it to show that the sort of freedom that is required for moral responsibility does not require alternative possibilities. Now, just permit the association of free will with that freedom and voilà, we have FW without AP.
The Direct Argument also requires near cousins of the principles of the fixity of the past and the laws (FP and FL). Instead of claiming that a person is powerless to render false the past and the laws, they claim nonresponsibility for the past and the laws. Let us label them FNP and FNL.

Having explained the details of the Consequence Argument, it is now easy to set out the Direct Argument since similar bits of exposition apply. Consider the following argument meant to show that if determinism is true at W, when George Bush gets down (s), he is not morally responsible for doing so:

1. \( \Box (P \land L \supset s) \) Assume determinism; apply it to s.
2. \( \Box (P \supset (L \supset s)) \) 1, and Logic.
3. \( \neg R (P \supset (L \supset s)) \) 2, Rule A.
4. \( \neg R P \) Assume FNP.
5. \( \neg R (L \supset s) \) 3, 4, Rule B.
6. \( \neg R L \) Assume FNL.
7. \( \neg Rs \) 5, 6, Rule B.

Call this argument Determined Bush. Points analogous to those pertaining to the Consequence Argument should apply here as well. For George Bush to be at least partly morally responsible for s at the determined world W, either he must be morally responsible for a fact of the remote past, or he must be morally responsible for a law of nature, or Rule B must be invalid. To emphasize the dialectical purpose of this argument, note that the conclusion, \( \neg Rs \), does not require that one “pass through” the further steps of employing FW and AP to reach the result that determinism is incompatible with moral responsibility; the result is direct.

In van Inwagen’s own estimation, the Direct Argument’s most vulnerable point turns upon its use of Rule B. He (1983, 188) writes:

If the compatibilist wishes to refute this argument . . . here is what he will have to do: he will have to produce some set of propositions intuitively more plausible than the validity of (B) and show that these propositions entail the compatibility of moral responsibility and determinism, or else

3. The reader might want a context that shows that George Bush’s getting down has the sort of moral significance pertinent to a judgment of responsibility (praise or blame). So suppose that he has promised his children that at a precise time he will get down in front of some television cameras in an effort to prove to them that he is “hip.” Were he to follow through, he’d be praiseworthy for keeping his promise.

354
Saying Good-bye to the Direct Argument the Right Way

he will have to devise a counter-example to (B), a counter-example that can be evaluated independently of the question whether moral responsibility and determinism are compatible.

John Martin Fischer and Mark Ravizza, and also David Widerker, opted for the strategy of producing counterexamples to Rule B. As my discussion unfolds, I will consider their efforts. But as will become clear, what should count as a counterexample depends in part on how to understand the dialectical force of the Direct Argument.

2. Unavoidability, the Direct Argument, and Widerker’s Farewell

When introducing the Direct Argument in *An Essay on Free Will*, van Inwagen began by reflecting upon potential objections to the Consequence Argument. In doing so, he considered a compatibilist prepared to resist his indirect argument by holding to some notion of avoidability different from the one featured in the Consequence Argument. Hence, as the imagined compatibilist objection goes, there is an equivocation between different senses of “could have done otherwise.” Therefore, a credible compatibilism has not been demonstrably refuted by any argument like the Consequence Argument. At a determined world, there is still a reasonable case to be made that an agent has the freedom to do otherwise. (Expressed in the terms set out in the previous section, determinism is compatible with AP.) To this interlocutor, van Inwagen (1983, 183) then writes:

I shall attempt to meet this charge of possible equivocation in another way. I shall argue “directly” for the incompatibility of determinism and moral responsibility. That is, I shall argue that determinism is incompatible with moral responsibility, and my argument will make no mention whatever of free will: notions like could have done otherwise and had a choice about will play no part in the argument in the present section.

So van Inwagen means to write an incompatibilist check to the compatibilist who is unprepared to give up a compatibilist picture of avoidability, even in light of the Consequence Argument. Van Inwagen’s dialectical burden, then, as he has cast it, requires that he cannot implicitly presuppose unavoidability in the Direct Argument.

It is perhaps the most impressive insight in Widerker’s assault upon the Direct Argument that, as he sees it, the notion of unavoidability is buried in one of the argument’s premises—and according to Widerker, it’s not buried all that deep either. Widerker asks, why is it
so clear that a person is not responsible for any of the laws of nature? He (2002, 324) offers the following diagnosis: *they are unavoidable*. If a person were able to avoid the laws of nature, then it would be dubious to claim that he or she is not morally responsible for them. For it would turn out that he or she could have prevented them from obtaining. Hence, Widerker holds, van Inwagen does not make good on his own dialectical promise to discharge the Direct Argument without relying upon the contested notion of unavoidability.

Widerker also presses another intriguing objection, one meant to prove the invalidity of any carefully restricted variation on the Direct Argument. He speculates that one sort of counterexample to Rule B will also apply to any efforts to improve upon it. To develop the crucial counterexample, Widerker considers how a compatibilist would react to the dialectical situation. Since van Inwagen has “set aside” avoidability in arguing with the compatibilist, he has not foreclosed the compatibilist’s use of her notion of it. And so, Widerker explains, imagine a compatibilist who is presented with the premises that a person is nonresponsible for the laws of nature and the facts of the remote past, and that he or she is nonresponsible for determinism’s obtaining. Now imagine that, at a determined world, the person knowingly does something morally wrong and believes that said action could have been avoided. A compatibilist who wishes to hold on to the notion of avoidability would still want to say that the person could have done otherwise in acting as he or she did. Hence, Widerker reasons, this compatibilist would be prepared to claim that the agent is morally responsible (that is, blameworthy) for the action: he or she knowingly did moral wrong, believing (correctly) that it was avoidable. Thus, a compatibilist would be able to appeal to his or her very own compatibilist notion of avoidability in the production of a counterexample to Rule B (Widerker 2002, 322–23). Given these considerations, Widerker argues, a proponent of the Direct Argument must first block a compatibilist’s appeal to avoidability. Therefore, the Direct Argument cannot be discharged independently of settling this battle.

3. A Reply to Widerker

Is Widerker correct to bid farewell to the Direct Argument in light of the two considerations mentioned above? This depends on how central one

4. The three Widerker mentions are ones formulated by Carl Ginet (2003), Eleonore Stump (2002), and me (2001).
Saying Good-bye to the Direct Argument the Right Way

takes the notion of avoidability to be in accounting for moral responsibility and for the sort of freedom or control pertinent to it. Clearly, Widerker invests a great deal in it. For instance, in imagining why one should accept the premise that no one is responsible for the laws of nature, Widerker (2002, 324) asks, “Why should one accept it if not for the fact that no one could have prevented them from obtaining?” I will return to this matter shortly. I turn first to his effort to prove the Direct Argument invalid with a counterexample to Rule B.

3.1. A Reply to Widerker’s Counterexample to Rule B

Is avoidability, when conjoined with knowingly doing moral wrong accompanied by the true belief of its avoidability, sufficient for moral responsibility (for either praise or blame)? Widerker thinks so, and he (2002, 323) offers the principle M, which he claims is “explicative of the very notion of blameworthiness”:

(M) Someone who knowingly acted in a morally wrong way, and believed correctly that he could have avoided acting as he did, is blameworthy for his act.

It is by appeal to (M) that Widerker takes himself to be entitled to a case in which the compatibilist can claim that the conclusion of the Direct Argument is false but all of its premises are true.

But M is false. The conditions that it specifies as sufficient for moral responsibility are not sufficient, at least not by the light of reasonable compatibilist standards. In particular, what is lacking is some requirement that when an agent commits a morally wrong action it is brought about in the “right sort of way.”5 Compatibilist accounts of “could have done otherwise,” such as the classical conditional account first suggested by Hume in *An Enquiry Concerning Human Understanding* and then advocated by Ayer (1954), Moore (1912), Schlick (1966),

5. On one account, this might require identification with the springs of one’s action (Frankfurt 1971). On another, it might require that the action arise as an appropriate response to a reason (Fischer and Ravizza, 1998; or Hajj 1998). On another it might require that an agent act from a suitably structured intention that allows for the possibility of strength (or weakness) of will (Mele 1995). On another, it might require that it arise from agential resources that include the general capacity to regulate behavior in accord with moral considerations (Wallace 1994). Or, finally, it might simply require the classical compatibilist demand that an agent act in the absence of coercion or compulsion so that the agent can be said to do what he or she wants unencumbered (Ayer 1954).
Davidson (1973), and others, are misguided if they are taken to exhaust the freedom-relevant condition for moral responsibility. Any account developed is better taken as an analysis of one necessary condition (a could have done otherwise condition) for a complete account of said freedom. Put in terms of the FW Condition and the AP Condition, in endorsing (M), Widerker is wrongly assuming that the FW Condition is exhausted in the AP Condition.

To illustrate what is lacking in Widerker’s M, imagine a case in which Adelle’s acting as she does at a determined world is counterfactually dependent upon her motives being thus and so. If they were otherwise, then she would act otherwise (just as the classical compatibilist’s conditional analysis required). Suppose that the action that she performs is throttling her daughter Danielle. Suppose also that she knows that doing this is morally wrong, and that she believes correctly (on the basis of compatibilist standards) that she could have done otherwise. But now imagine that the motive causing her action is the upshot of an insane sensitivity to the reasons for so acting. Adelle, as it turns out, suffers from paranoid schizophrenia. She believes that Martians speak to her. These space aliens bid her to do terrible things, which she does know are morally wrong, but which provide her with strong but not irresistible desires to do these things. When Adelle throttles Danielle, knowing that it is morally wrong to do so, her motive for her action is that she believes that terrorizing her daughter will appease the Martians. Furthermore, Adelle would do otherwise and not throttle Danielle, but only for considerations that would prove that she is not appropriately responsive to reasons. For instance, Adelle would not attempt to appease the Martians by throttling Danielle if she came to believe that the Martians would rather have her take Danielle out for coffee than throttle her, or if other space aliens, Saturnians, were instructing her not to do the Martians’ bidding. In such a case, everything that is required for Widerker’s (M) is satisfied, and yet Adelle is not blameworthy for throttling Danielle. What explains Adelle’s not being blameworthy is not that (by compatibilist standards) she could not do otherwise. What explains why she is not blameworthy concerns the character of the actual causal history of her action.6

In reply to Widerker, van Inwagen can exploit these further requirements for moral responsibility (and blameworthiness). Thus van

---

6. The case of Adelle and Danielle exploits Fischer and Ravizza’s source conditions for moral responsibility. What Adelle lacks is what Fischer and Ravizza call a proper receptivity to reasons. (See Fischer and Ravizza 1998, 65–72.)
Saying Good-bye to the Direct Argument the Right Way

Inwagen can claim that the Direct Argument applies its force to claims of moral responsibility irrespective of considerations of avoidability. Under the assumption of determinism, nonresponsibility for the facts and the laws transfers to nonresponsibility for conduct due to factors necessary for moral responsibility other than avoidability. If so, then the Direct Argument is not threatened by allowing avoidability.

3.2. A Reply to Widerker’s Challenge: Explaining Nonresponsibility for the Laws

Now consider Widerker’s provocative question, “Why should one accept it [the premise that no one is even partly morally responsible for the laws of nature] if not for the fact that no one could have prevented them [the laws of nature] from obtaining?” (Widerker 2002, 324; my brackets). What exactly does Widerker mean to convey in asking this question? He implies that unavoidability of the laws is a sufficient condition of nonresponsibility for them. This itself is subject to dispute. There is a highly contested debate regarding the claim that unavoidability is sufficient for lack of morally responsibility. Of course, Frankfurt (1969), along with many others, argues that this is false.7 But for now, I shall put Frankfurt’s thesis aside. Grant Widerker the contested assumption. What more is there to his remark about nonresponsibility for the laws?

Widerker also seems to suggest that there is no sufficient condition of nonresponsibility for the laws other than unavoidability. Should we therefore infer that, as Widerker sees it, unavoidability of the laws is also a necessary condition of nonresponsibility for them? In my estimation, this interpretation is too uncharitable. It amounts to the highly implausible thesis that avoidability of the laws is by itself a sufficient condition of responsibility for them. But surely conditions sufficient for moral responsibility (for any x) are not exhausted by appeal to avoidability alone. That thesis is a nonstarter. I think that the best way to interpret Widerker’s remark about the laws of nature is as follows: Widerker is asking an opponent like van Inwagen, who asserts nonresponsibility for the laws, to demonstrate the assertion in the absence of the one (allegedly) clear basis for establishing it—one relying on unavoidability.

Widerker (2002, 324) considers and rejects one plausible proposal, “that no human being brought them about.” Why, according to Widerker, will this not do? Because moral responsibility for a fact, he

7. In unpublished work, Charles Cross has developed this point.

359
 MICHAEL MCKENNA

maintains, does not require causal responsibility for it. Widerker cites the case of an omission in which a person’s failure to act knowingly allows a bad consequence to unfold. In this case, one person allows another to be bitten by a cobra even though the first person knew the event was about to take place and could have prevented it. Is Widerker’s remark here enough to dismiss the claim that nonresponsibility for the laws can be established by showing that no one brings them about? I do not think so. Granted, there are cases in which, on certain views, causal nonresponsibility is not sufficient for moral nonresponsibility. But there are also cases in which it is. For example, I am not morally responsible for being five feet, ten inches tall because I did nothing to bring about my height. Even supposing that in the case of omissions involving consequences, causal nonresponsibility does not yield moral nonresponsibility, this does not show that causal nonresponsibility for the laws of nature is not sufficient to account for moral nonresponsibility.

Reflect on the claim that no one is morally responsible for the laws of nature and assume that the laws of nature are both unavoidable and not brought about by any human agent (which I assume is the truth of the matter.) It seems especially relevant in accounting for nonresponsibility for the laws of nature that no one brings them about. This, and not their unavoidability, seems to have a greater explanatory bearing upon our preparedness to accept FNL. In fact, Widerker understates his opponent’s position in suggesting that FNL could be based upon the fact that no one brought about a law of nature. True, no one brought about a law of nature. But there is also a stronger modalized claim to be made here. It is also true that no one is able to bring about a law of nature. This includes any law of nature, not just the ones that do obtain. A law of nature is understood, at least in part, in terms of limitations upon us. If the laws of nature set for us the limits of our abilities (and they do), then surely we are not able to bring them about—any of them. Beyond this point, there is yet a further point one might make regarding

8. Perhaps on a Humean version of the laws, this is false. But I set this proposal aside. For a discussion of how a Humean account would affect the Consequence Argument, see Beebee and Mele 2002.

9. This point is about the relationship between the laws and our abilities. It is not meant to be generalized. Certainly it is possible for one to act so as to limit one’s own abilities. A person can take a drug that renders him or her incapable of walking. The difference is that his or her ability to walk was not impaired prior to taking the drug. But the manner in which the laws set limits upon our abilities differs; the laws of nature apply throughout the entire history of a person’s life. I thank Al Mele for pressing me on this matter.
Saying Good-bye to the Direct Argument the Right Way

the actual basis for the laws of nature. Ted Warfield (1996, 218) has written that the truth-makers for the laws of nature “obtained prior to the existence of any human agents.” Warfield (ibid., 224n4) maintains that this feature of the laws is at least as compelling a basis for nonresponsibility for them as is any consideration about lack of human control over the laws.10

In summary, were one to assess the basis for claiming that no one is morally responsible for the laws of nature, there are three distinct considerations one might appeal to: unavoidability, the inability to bring any of them about, and the fact that their truth-makers arose in the absence of the existence of any human agents at all. In my estimation, Widerker has not said enough to deny that nonresponsibility for the laws of nature can be explained by appeal to the fact that no one produces them. To the contrary, the fact that no one produces the laws of nature—indeed, that no one can produce any law of nature—provides a powerful resource for the claim that no one is morally responsible for them. The resources are even further supplemented when one considers Warfield’s point that the truth-makers for the laws of nature were settled prior to the existence of any human agents. Hence, it seems that van Inwagen has adequate resources available to account for nonresponsibility for the laws besides considerations regarding unavoidability.

3.3. A Reply to Widerker’s Assumption Regarding Unavoidability

I conclude that Widerker’s farewell to the Direct Argument rests on an estimation of the role of avoidability in accounting for responsibility that is very strong, stronger than is warranted. Therefore, it is not clear that the compatibilist who feels entitled to his or her preferred notion of avoidability can appeal to it to produce counterexamples to Rule B or to claim that the premise of nonresponsibility for the laws rests on the assumption that the laws are unavoidable. Hence, van Inwagen does not seem to have violated his promise to produce an argument that does not rest upon unavoidability.

In closing this section, I predict that Widerker has not said his final farewell to the Direct Argument. First, supposing, as van Inwagen does, that the dialectical purpose of the argument is to achieve its result

10. Of course, Warfield’s conception of a law must be one in which facts about causal structure are settled by an early time in the history of a deterministic world. Though contestable on Humean grounds, I agree.
irrespective of importing the notion of unavoidability, I have argued that
Widerker is not clearly correct to claim that it fails on this front. Sec-
ond, in the following section I will question whether the effort to do
without the assumption of unavoidability is the most constructive dialec-
tical purpose to which the Direct Argument should be put. If it is not—
if there is no dialectical burden on the proponent of the Direct Argu-
ment to get by without presupposing the incompatibility of determinism
and avoidability—then Widerker’s objections to the Direct Argument
are otiose. Hence, he is not saying good-bye to it in the right way.

4. Source Incompatibilism, and the Problem with Fischer’s Good-bye

There are more dialectical moves that need to be sorted out. Both van
Inwagen and Widerker share the assumption that the dialectical force of
the Direct Argument is meant to speak (at least in part) to compatibilists
who retain a commitment to the compatibility of avoidability and deter-
mindism. But there are various theorists who see the dialectical force of
the Direct Argument quite differently. In particular, there are both com-
patibilists and incompatibilists who agree that determinism rules out
avoidability, but who are convinced by other arguments—almost exclu-
sively, by some variation on Frankfurt’s argument (1969)—that the sort
of freedom pertinent to moral responsibility does not require avoid-
ability (Fischer 1994; Fischer and Ravizza 1998; Haji 1998; McKenna
2003; Mele 1995; Stump 1996; Pereboom 2001; and Zagzebski 2000).
These source theorists, and especially the source compatibilists, would not
think themselves entitled to Widerker’s (alleged) counterexample set
out above. They would not think themselves entitled to a case in which a
person at a determined world is not responsible for the facts of the past
and the laws of nature, and is not responsible for their joint entailment
of one possible future, but is responsible for his or her action because he
or she could have avoided acting as he or she did. These philosophers
would not think themselves entitled to the “because” that is at the core
of Widerker’s proposed counterexample to Rule B.

Source theorists agree that whatever accounts for moral responsi-
bility and the sort of control or freedom that grounds it is to be found
in what an agent does and his or her basis for doing it, not in
what else he or she could have done. Rejecting avoidability as a crucial
necessary condition (the AP condition set out in section 1 above), the
debate between source compatibilists and source incompatibilists turns
on whether determinism precludes fulfillment of the proper source
Saying Good-bye to the Direct Argument the Right Way

conditions for freedom and responsibility. The appeal of the Direct Argument to the transfer of nonresponsibility for facts before one’s birth and the laws of nature through to the actions one performs is one incompatibilist effort to show that determinism does in fact preclude fulfillment of these conditions.11 It is for this reason that I agree with Fischer (forthcoming) that we source compatibilists should not bid the argument farewell for the sorts of reasons of avoidability at the heart of Widerker’s farewell. Rather, source compatibilists should greet the argument with a hello since it speaks directly to the source compatibilist concerning the requirements on sourcehood for moral responsibility. Though in my estimation, and clearly in Fischer’s too, it had better be a very guarded hello. The Direct Argument is a serious challenge to source compatibilists, and it is only proper for source compatibilists to consider it carefully. But, as with Fischer, after considering the Direct Argument within the dialectical space of the source theorist, I too eventually want to send it packing. However, my reasons differ from Fischer’s.

Why is it that Fischer wants to reject the Direct Argument? In his coauthored work with Mark Ravizza, Fischer, along with Widerker, accepts van Inwagen’s challenge to produce a counterexample to Rule B, which Fischer and others refer to as the Principle of the Transfer of Nonresponsibility, or just Transfer NR. (Henceforth, I’ll shift to speaking in terms of Transfer NR.) Fischer and Ravizza produce a series of “two-path” cases in which one of the paths allows a person to be free and morally responsible for something: an act, an omission, or the consequences of an act or omission. But there is another path that the agent is not responsible for that also inevitably leads to the same something. So here is the example of theirs, Erosion, that has received the most attention:

Imagine that Betty plants her explosives in the crevices of a glacier and detonates the charge at T1, causing an avalanche that crushes the enemy fortress at T3. Unbeknownst to Betty and her commanding officers, however, the glacier is gradually melting, shifting, and eroding. Had Betty not placed the dynamite in the crevices, some ice and rocks would have broken free and crushed the enemy base camp at T3. (Ravizza 1994, 72–73)

The case is meant to be a counterexample to Transfer NR because Betty is not responsible for the facts of the erosion or for the fact that it will inevitably lead to the destruction of the camp, yet she is responsible for the destruction of the camp. This example first appeared in Ravizza’s earlier publication (1994) but then figured centrally in Fischer and Ravizza’s joint contributions. The example is a case of responsibility for consequences, and some have rejected it for that reason, claiming that it does not get to the heart of the debate (Pereboom [2001]; Widerker [2002]). Restricting Transfer NR just to cases of simple mental actions—to the loci of freedom and responsibility—would generate different results, or so the objection goes. But then it is open to Fischer and Ravizza to attempt the same argument as applied to simple mental actions. Thus, I have objected to their assault on the Direct Argument for different reasons.

In earlier work (2001), I replied to Fischer and Ravizza on behalf of the source incompatibilist by granting that, strictly speaking, Erosion is a counterexample to Transfer NR.12 But, I argued, Erosion would not be a counterexample to a suitably revised version of Transfer NR, Transfer \( NR^* \), that was restricted to one-path cases.13 Here is a slightly amended version of the formulation I offered (2001, 45):

1. \( p \) at time \( T_1 \), and no one is or ever has been even partly responsible for the fact that \( p \);
2. i. \( p \) is part of the actual sequence of events \( e \) that gives rise to \( q \) at \( T_2 \) (where \( T_2 \) is later than \( T_1 \));
   ii. \( p \) is causally sufficient for the obtaining of \( q \) at \( T_2 \), and any other part of \( e \) that is causally sufficient for \( q \) either causes or is caused by \( p \);
   iii. no one is or ever has been even partly responsible for \( 2.i. \) and \( 2.ii. \);
3. Therefore, no one is or ever has been even partly responsible for the fact that \( q \) obtains at \( T_2 \).

Informally, the idea is to be captured in this way: if a person is not responsible for a fact, and if this fact inevitably leads to another, and if,
furthermore, there is no other path that leads to the latter fact and the person is not responsible for the path leading to the latter fact, then that person is not responsible for the resultant fact either. Erosion is no counterexample to Transfer NR**, nor could any two-path case defeat it since it is restricted to one-path cases. This means that, if it is a reasonable restriction, Fischer and Ravizza cannot say their final good-byes to the Direct Argument either.

Is my effort to restrict Transfer NR** to one-path cases defensible? I’ll get to that. But let me make clear the intuitive basis for my criticism of Fischer and Ravizza’s assault on the Direct Argument. My central objection is that the manner in which the source incompatibilist takes determinism to rule out moral responsibility and the manner in which the Direct Argument can be employed on behalf of the source incompatibilist is not relevantly like a two-path case of the sort Fischer and Ravizza enlist to reject Transfer NR. If determinism is true, it is not as if there is one path to a certain result in virtue of which determinism is true, and then there is some other distinct path, unhinged from the former, that a morally responsible agent initiates so that he or she bears responsibility for the object of responsibility in question (be it an act, an omission, or a consequence). The facts settled by the truth of determinism include the facts pertaining to the “action path” in virtue of which the agent is alleged to be responsible.14

5. Fischer’s Reply: Why He Thinks He Can Say Good-bye

In his characteristically thoughtful manner, Fischer replied to my restricted one-path version of Transfer NR. I will offer a counterreply to his two core points against me. As to the first, Fischer (2004, 200–201) writes:

I concede McKenna’s point about constitution, but I do not thereby accept that moral responsibility is necessarily crowded out. A one-path

14. For a similar point, though put to a different use, see Pereboom’s indictment of libertarian efforts to appeal to nonreductive materialism to save freedom. Pereboom (2001, 71) writes:

Thus, even if macrophysical indeterminism is true, as long as the microphysical level is governed by deterministic laws, all of our decisions will be rendered inevitable by virtue of previous states of the universe, just as their microphysical realizations are. All of our decisions will then be alien-deterministic events—events for which factors beyond the agent’s control determine their occurrence—and none of them could be agent-caused free decisions.
case could be thought of as two separate paths that are superimposed. Alternatively, I would recommend that we think of the one path as having (at least) two separate sets of features. Suppose that Betty deliberates just as she actually does in the Erosion cases, but there is no “second path”—no glacier that will independently cause the upshot (the destruction of the enemy camp at T3). But imagine further that causal determinism obtains in the actual path that leads to Betty’s decision and action. One set of features entails that there is along this path just the sort of unimpaired deliberation that obtains in Erosion...—practical reasoning and subsequent action undistorted by factors that uncontroversially threaten moral responsibility. But of course another set of features renders the actual sequence of events causally deterministic, and causal determinism is alleged by the incompatibilist to rule out moral responsibility.

Note, however, that causal determinism is not a factor that uncontroversially rules out moral responsibility. The features in virtue of which the actual path is causally deterministic are thus importantly different from such a feature as an avalanche caused by movements of a glacier, which uncontroversially fails to confer responsibility.

I am especially attracted to Fischer’s reply to me in this passage. My own convictions in other arenas of metaphysics are consistent with accounts of the constitution of material objects according to which objects like statues, toasters, and Homo sapiens are not identical with the physical material out of which they are constituted. Hence, why not a single event-pathway constituted by multiple sets of distinct, causally sufficient conditions? If this is possible, and if one of them (at worst) only controversially threatens satisfaction of pertinent freedom or control conditions, then why not concede that Fischer and Ravizza can press their reply to the Direct Argument even if they do restrict their reply to one-path cases?

Maybe I should. But three considerations cast doubt on Fischer’s proposal. First, constitution might be identity. Second, it is unclear that we can speak of two distinct paths that are superimposed in the cases that Fischer has in mind. If all there is involves just one single deterministic path, then it might rule out moral responsibility after all. Given Fischer’s particular strategy, what is required for him to make sense of the claim that we can imagine two distinct paths superimposed on a single event-pathway is that the features that involve the responsibility-conferring characteristics are distinct from the features in virtue of which determinism obtains. One might object that it would not be as if there were other properties that were the ones in virtue of which determinism obtains and then, distinct from these, the freedom- and
Saying Good-bye to the Direct Argument the Right Way

responsibility-conferring properties. Third, even granting that Fischer can make good sense of the idea of two paths superimposed on a single event-pathway, might it not also be the case that the “upper-level” action path is also determined, distinct from the “lower-level” path (presumably characterized in terms of microstructure)? If so, the proponent of the Direct Argument will hold that by virtue of determinism, nonresponsibility is transferred even along this upper-level path.15

While I sympathize with Fischer’s reply to me regarding the possibility of restricting Transfer NR to one-path cases, in my estimation, for the reasons set out in the previous paragraph, this proposal is too controversial to rest a rejection of the Direct Argument on it. But now consider another of Fischer’s replies:

I claim then that the attempt to employ Transfer NR, restricted to one-path causally deterministic contexts, as part of an argument for the incompatibility of causal determinism and moral responsibility, issues in a dialectical stalemate. This implies that such an attempt does not generate a successful argument for incompatibilism. . . The argument that employs some restriction or qualification of Transfer NR employs a principle that can only be established by begging the question against the compatibilist; it does not employ a principle that is broadly appealing. (Fischer 2004, 203)

Here Fischer makes two claims. First, all restrictions to one-path cases must carry the modifier “deterministic.” Second, no credible sort of restriction to one-path cases could be “broadly appealing”; rather, any such restriction is question begging. I take this second point to mean that no restricted principle could be shown to have a legitimate intuitive basis independently of incompatibilist commitments. I’ll take each in turn.

First, Fischer is correct that, as I cast my variation of Transfer NR, it was restricted to cases involving causally deterministic relations, as is made clear by speaking in terms of “causally sufficient conditions.” However, the principle could be recast so as to weaken the causal relation in a way that allows for probabilistic causes. Still, it would preserve the intuitive idea of transferring nonresponsibility along causal chains in which an agent is not responsible for the end product of the chains so long as he or she is not responsible for any of the causally contributing factors along the way. Informally, the idea is to be captured in this way:

15. I owe this point to an anonymous referee for Philosophical Review.
Transfer**: If a person is not responsible for a fact, and this fact is at some particular time the sole causal antecedent of the path to the bringing about of a further fact, and if the person is not responsible for any part of the path between the two facts, then he or she is not responsible for the further fact either.\(^{16}\)

Transfer** would do the requisite work for the proponent of the Direct Argument even if the “bringing about” relation is weaker than a causally deterministic one.\(^{17}\) Thus it would not warrant Fischer’s objection that it is restricted to deterministic contexts.

Second, is this revised version of Transfer NR question begging? If it is, it is not clear from what Fischer has said why it is question begging. It seems completely neutral between compatibilists and incompatibilists. That it commands an unbiased, intuitive appeal seems confirmed by the sorts of cases that van Inwagen himself offers to establish the credibility of his Rule B. He mentions the following two cases, which we can simply label Snakebite and Plato:

\[
\text{NR John was bitten by a cobra on his thirtieth birthday;}
\]
\[
\text{NR (John was bitten by a cobra on his thirtieth birthday } \supset \text{ John died on his thirtieth birthday);}
\]
\[
\text{NR John died on his thirtieth birthday;}
\]
\[
\text{and:}
\]
\[
\text{NR Plato died in antiquity;}
\]
\[
\text{NR (Plato died in antiquity } \supset \text{ Plato never met Hume);}
\]
\[
\text{NR Plato never met Hume.}
\]

Simply amend these cases so that they include the qualification that there is no further “path” that did realize either the fact that John died on his thirtieth birthday or that Plato never met Hume. Don’t these cases invite a preparedness to accept some restricted one-path version

\(^{16}\) I am indebted to one of the referees for Philosophical Review for suggesting this wording.

\(^{17}\) This principle is similar in spirit to one Pereboom advances when arguing that nondeterministic causal sources are also freedom and responsibility undermining in the absence of some sort of control. Consider, for example:

We have already seen that by incompatibilist intuitions, an agent is not responsible for decisions determined by factors beyond his control. However, if these factors, rather than determining a single decision, simply leave open more than one possibility, and the agent plays no further role in determining which possibility is realized, then we have no more reason to hold him responsible than we do in the deterministic case. (Pereboom 2001, 48)
Saying Good-bye to the Direct Argument the Right Way

of Transfer NR? Imagine the many more cases one can generate. Here is just one, Water Pipes:

NR The temperature is dropping rapidly;
NR (The temperature is dropping rapidly ⊃ Mrs. Weatherby’s water pipes burst);
NR Mrs. Weatherby’s water pipes burst.

Now also amend this case so that it includes the qualification that there is no further path that did bring it about that Mrs. Weatherby’s water pipes burst. In my earlier piece (2001, 51n45), I pointed out that the credibility of a version of Transfer NR restricted to one-path cases is confirmed by the very fact I am bringing into relief here: cases like van Inwagen’s Snakebite and Plato, as well as my present case of Water Pipes, can just as easily be accommodated by a restricted version. Fischer concedes that the cases can be accommodated by my proposed rendition of Transfer, but he maintains that this is not the issue. He claims that there is yet a further rendition that accommodates these uncontroversial cases, but that will not allow the incompatibilist to have his or her way with the principle in the service of the Direct Argument. In his proposed rendition, Fischer (2004, 202) offers the following premise as the one crucially different from the one I offered:

On the actual path that leads from p’s obtaining to q’s obtaining, either there is no factor that at least prima facie could be thought to ground moral responsibility, or there is some factor that uncontroversially undermines moral responsibility.

As far as I can tell, a version of Transfer NR amended to incorporate this premise would be question begging, but it would be question begging in favor of a compatibilist construal. Why? Because this formulation in effect ensures that the principle of transfer of nonresponsibility cannot be used against the compatibilist.

It looks as if we have reached a dialectical stalemate. Fischer charges that my incompatibilist-friendly amendment to Transfer NR begs the question against the compatibilist since it guarantees incompatibilist results. But his compatibilist-friendly rendition begs the question against incompatibilists since it prematurely forecloses an application for incompatibilist conclusions. It is precisely at this point that my compatibilist response to the Direct Argument differs from Fischer’s. I do not think that the proper compatibilist response is to offer a competitor version of Transfer NR at all, a move that quickly invites a
countercharge of begging the question. Rather, the compatibilist should resist the inclination to postulate any such principle. In the following section, in an effort to vindicate Fischer’s suspicion about principles such as Transfer NR**, I will argue that relevant incompatibilist-friendly versions of Transfer NR are indeed suspect. But in doing so, I will not attempt to offer a compatibilist competitor as a replacement principle.

What I find missing in Fischer’s charge of begging the question is some explanation of why the incompatibilist-friendly renditions appear so reasonable. It is not just worries about incompatibilism that give force to these inference principles. It’s the very idea, as illustrated in cases like Snakebite, that causal paths of nonresponsibility seem to transfer so cleanly. My diagnosis of Fischer’s treatment of a principle like Transfer NR** or Transfer NR*** is that if indeed these inference principles are corrupt, how it is that they are corrupt needs to be brought into relief. What Fischer has offered thus far does not explain away their appearance of dialectical innocence as natural extrapolations from the sorts of cases enlisted above.

In brief, my overall counterreply to Fischer is that, if the prospects of making two-path cases converge on a single event-pathway is too controversial to hang the Direct Argument upon, then he has only the resources of one-path cases to work with. It is then imperative for his good-bye to the Direct Argument that any one-path formulation of Transfer NR is question begging. But I have argued that if it is, he needs to explain why it seems not to be. More needs to be said. Upon reflection, I think that Fischer is on to something, and I shall now develop a different tactic that will help bring into relief how Transfer NR should be undermined. In doing so, I shall set aside the question as to whether, as Fischer suggests, the incompatibilist is begging any questions in offering a formulation such as Transfer NR**. Instead, I shall explain away the initial appearance of credibility in any formulation of Transfer NR as it is alleged to arise out of reflection on cases such as Snakebite, Plato, or Water Pipes.

6. Why Say Hello to Transfer NR in the First Place?

How is it that any unrestricted version of Transfer NR became a viable candidate for the generation of an incompatibilist argument? I shall argue that it was dialectically inappropriate for van Inwagen to use
Transfer NR in an incompatibilist argument. The unfolding dialectic should never have even gotten off the ground.

Let me start with a transparent observation about the dialectic. Here is a candidate instance for Transfer NR that all contestants in the debate will agree is one that cannot be used in the service of arguing for the inference principle. Suppose that in the example in which George Bush gets down, he does so at a determined world, and that just prior to exercising his (putatively) free will in deciding to get down, he deliberates about doing so. He considers reasons against, such as the fun that comedians will poke at him for his effort, and reasons for, such as the promise he made to his daughters to do what he could to appear more hip. And suppose also that as a causally deterministic upshot of his deliberative process, he then chooses to get down and does so. Now consider the example, Bush Deliberates:

\[
\text{NR Bush deliberates about getting down } \bullet \ L; \\
\text{NR (Bush deliberates about getting down } \bullet \ L. \supset \text{ Bush gets down);} \\
\text{NR Bush gets down.}
\]

Clearly, Bush Deliberates could not be enlisted to establish Transfer NR in the context of this debate; it could only be asserted after Transfer NR is already established. For the dispute between compatibilists and incompatibilists turns crucially on the validity of arguments such as Bush Deliberates.

Now consider an unobjectionable instance of Transfer NR, but one that exploits a relation different from the typical sort. The typical sort exploits causally sufficient conditions. Instead, consider the case Solar System, which exploits a broadly logical relation.

\[
\text{NR The planets number less than eleven;} \\
\text{NR (The planets number less than eleven } \supset \text{ the planets number less than twelve);} \\
\text{NR The planets number less than twelve.}
\]

Consider now the proposition that the fact that the planets number less than eleven entails the fact that the planets number less than twelve:

\[\Box (\text{The planets number less than eleven } \supset \text{ the planets number less than twelve})\]

This proposition, along with Rule A, generates the second premise in the immediately preceding argument. What makes this entailment, \[\Box (\text{The planets number less than eleven } \supset \text{ the planets number less than twelve}),\] true? Plainly, some simple arithmetic rules, or more
immediately, the fact that, necessarily, eleven is less than twelve. That is all. My assault on the Direct Argument begins with thinking carefully about the facts in virtue of which any one of these crucial entailments would be made true. For this in turn will cast light on the facts in virtue of which the nonresponsibility material conditional claims, such as NR (The planets number less than eleven ⊃ the planets number less than twelve), are (allegedly) rendered true.

So consider van Inwagen’s case Snakebite. What are the fact-making characteristics of:

□ (John was bitten by a cobra on his thirtieth birthday • L. □ John died on his thirtieth birthday)?

I do not know nearly enough about snake bites, their venom, or human physiology to state with any clarity what facts make it true that getting bit by a cobra will (or can) kill a person. But I believe I know enough to know that, unlike the case Solar System, there is a discrete collection of facts that stand in relations of causal dependence to each other so that at one “far end” John gets bitten by a cobra, and at the other “far end,” John dies. Between these “ends” there is a series of causal relations that can be expressed as discrete causal facts. For instance, in the causal path from John’s getting bit to John’s dying, there is the fact that John’s getting bit caused it to be the case that cobra venom entered John’s bloodstream. The event of cobra venom entering John’s blood stream then caused the unfolding of yet a further causally connected chain of events all pertaining to John’s physiology, and this chain’s terminus then caused John’s death. To make the point clear, consider all of the following propositions:

\[ p = \text{John was bitten by a cobra on his thirtieth birthday;} \]
\[ r = \text{Cobra venom entered John’s bloodstream;} \]
\[ s = \text{a causal chain of events pertaining to John’s physiology unfolded;} \]
\[ q = \text{John died on his thirtieth birthday.} \]

If some particular arrangement of causal events make it true that \((p \cdot L. □ q)\), this is because the laws of nature being thus, along with the initial conditions identified as \(p\), entail that \(p\) causes \(r\); \(r\) causes \(s\); and \(s\) causes \(q\).

18. In An Essay on Free Will (1983, 187), van Inwagen omits the conjunct ‘L.’ in the antecedent of this conditional. But it is plausible to suppose that he assumed it.
So consider the following conjunction as it bears on van Inwagen’s case, Snakebite:

\[(S) \Box (p \land L. \supset r) \land \Box (r \land L. \supset s) \land \Box (s \land L. \supset q)\]

If \(\Box (p \land L. \supset q)\) is true, then so is (S) because (S) is what makes the former true. My contention is that in any candidate instance of Transfer NR, such as Snakebite, the relevant material conditional premise of non-responsibility, if it involves a claim of causal sufficiency, can be treated in like fashion. The pertinent conditional premise issues from an application of van Inwagen’s Rule A to some proposition such as the entailment \(\Box (p \land L. \supset q)\). And if this proposition is true, then so is some analogue to (S).

Given the above considerations, it is natural to think that the same point applies to the conjunction generated by applying van Inwagen’s Rule A to each of the conjuncts in (S):

\[(S') \text{NR} (p \land L. \supset r) \land \text{NR} (r \land L. \supset s) \land \text{NR} (s \land L. \supset q)\]

Given (S’) we can construct a chain of arguments built from each of the conjuncts:

1. NR (p \land L) given
2. NR (p \land L. \supset r) (S’)
3. NRr 1, 2, TNR
4. NR (r \land L) 1, 3
5. NR (r \land L. \supset s) (S’)
6. NRs 4, 5, TNR
7. NR (s \land L) 1, 6
8. NR (s \land L. \supset q) (S’)
9. NRq 7, 8, TNR[^19]

Now suppose any one of the links in the chain of arguments generated from (S’) was subject to credible doubt. For instance, suppose that someone was partially responsible for the fact that, after getting bit by a cobra, the venom entered John’s blood stream (maybe a nurse was nearby who could have prevented this but did not). Then steps 1 through 3 would be in dispute. If any link were subject to credible doubt, then that status would impugn the argument Snakebite as well. In the case of Snakebite, given

[^19]: This particular formulation of my argument was kindly suggested by an anonymous referee for *Philosophical Review*. I am also grateful to Carl Ginet for help in developing this line of reasoning.
natural assumptions about the way a poisonous snakebite works, there is just no reason to think that any of the links is dubious. Hence, there is no reason to think that Snakebite is questionable. Were we to work instead with the case of Plato or Water Pipes, we would get the same result. We will not, however, if we work with a very different sort of case.

Recall the original case, Determined Bush, used in section 1 to illustrate an application of the Direct Argument. To bring into relief the relevant application of Transfer NR as it figured in the argument Determined Bush, I shall summarize it as follows:

\[ \text{NR (P \& L)} \]
\[ \text{NR (P \& L. \supset s)} \]
\[ \text{NRs} \]

Just as we asked about the facts in virtue of which the crucial entailment is true (if it is) in the case of Snakebite, so too can we ask about the facts in virtue of which the proposition \( \Box (P \& L. \supset s) \) is true (if it is). This proposition, along with Rule A, is what generates the second premise in the immediately preceding argument. In answering this question, I shall merge Determined Bush with Bush Deliberates. Call this case, Determined Bush Deliberates. Consider then all of the following ingredients: ’P’ in the proposition \( \Box (P \& L. \supset s) \) is meant to obtain at a time in the remote past, a time long before the birth of George Bush. We can assume that after that time, \( g: \) some very large constellation of discrete events in a causal chain led to the existence of the person George Bush and also his subsequent acquisition of the deliberative capacities that he employed at the very moment he deliberated about getting down. Let us designate the proposition that Bush deliberates about getting down as \( r \). Finally, recall that the proposition that Bush gets down at a certain time is identified as \( s \). In summary, we have:

\[ P = \text{The facts long before George Bush was born were thus and so;} \]
\[ g = \text{a causal chain led to the existence of Bush and his deliberative capacities;} \]
\[ r = \text{Bush deliberates about getting down;} \]
\[ s = \text{Bush gets down at a certain time.} \]

We now have the makings of the following treatment of the pertinent entailment \( \Box (P \& L. \supset s) \) found in Determined Bush as it figures in the expanded case Determined Bush Deliberates:

\[ (DBD) \Box (P \& L. \supset g) \bullet \Box (g \& L. \supset r) \bullet \Box (r \& L. \supset s) \]
Applying van Inwagen’s Rule A to each of the conjuncts of (DBD), we get:

$$(DBD') \text{NR} (P \cdot L. \supset g) \cdot \text{NR} (g \cdot L. \supset r) \cdot \text{NR} (r \cdot L. \supset s)$$

Just as we did with (S'), so too, given (BD') we can construct a chain of arguments built from each of the conjuncts:

1. NR $(P \cdot L)$ given
2. NR $(p \cdot L. \supset g)$ (DBD')
3. NR $(g \cdot L)$ 1, 2, TNR
4. NR $(g \cdot L. \supset r)$ (DBD')
5. NR $(r \cdot L)$ 4, 5, TNR
6. NR $(r \cdot L. \supset s)$ (DBD')
7. NR $r$ 7, 8, TNR
8. NR $(r \cdot L. \supset s)$ (DBD')
9. NR $(r \cdot L. \supset s)$ (DBD')

This is the argument Determined Bush Deliberates.

Determined Bush Deliberates yields a highly undesirable result for the proponent of the Direct Argument. The last link in the chain of arguments above, the one involving steps 7 through 9, just is Bush Deliberates, the very one that figures in what earlier I stated could not be employed in support of Transfer NR. Such an instance, I claimed, could not be used in the service of establishing Transfer NR. Even if it is determined that Bush comes to deliberate as he does, and even if he is not responsible for the causal history that gives rise to his now being in that position, it does not follow uncontroversially that he is not responsible for the relationship between his states of deliberation and his actions, nor for his actions themselves—given the dialectical situation between compatibilists and incompatibilists. Furthermore, just as I argued above, if one link in the chain of arguments is subject to credible doubt, as it is in Determined Bush Deliberates, then that impugns the argument Determined Bush as well.

Naturally, the proponent of the Direct Argument will object that I have not set up the dialectic fairly. For I originally stated in introducing

---

20. It should be noted that this point is one available to either the sort of compatibilist who requires alternative possibilities (the sort Widerker considered in his assault upon the Direct Argument), or a source compatibilist, who denies the relevance of alternative possibilities.
the case Bush Deliberates that it could not be employed to argue for the inference principle. It could only be assumed to be a legitimate inference after the principle was established. So why not treat cases like Snakebite, Plato, and Water Pipes as establishing the validity of Transfer NR, and then apply it to cases like Determined Bush, Bush Deliberates, and Determined Bush Deliberates? The reason is this, which is crucial to my argument: In Snakebite, Plato, and Water Pipes, the chains of causal sufficiency through which any transfer of nonresponsibility is transmitted never “pass through” a normally functioning agent who exercises unimpaired deliberative capacities in the production of an (allegedly) free action for which he or she is morally responsible. If the only uncontroversial cases one can cite to establish Transfer NR are such cases, this strongly indicates that Transfer NR is best restricted to cases that do not “pass through” such agents.21

The catch for advocates of the Direct Argument is as follows. They must argue for Transfer NR by introducing an example that differs from Snakebite, Plato, or Water Pipes in the requisite way—that is, they must produce a valid example that aids in confirming Transfer NR by including a causal chain passing through a normally functioning agent in the needed way.22 Hence, at this stage in the dialectic, they will have to introduce some proposition like the one featured in Determined Bush, NR (P • L. ⊃ s). But I have argued that any such proposition is true only if a proposition such as (DBD’) is, and this will yield chains of arguments like Determined Bush Deliberates that have embedded in them highly controversial arguments such as Bush Deliberates.23

21. To borrow Slote’s manner of expression, we have very good reason to believe that the necessity expressed in claims of nonresponsibility is “selective” in the inferences it licenses.

22. Bear in mind that van Inwagen argued for Transfer NR (his Rule B) by using confirming examples of it. If the examples he has offered cannot do the work of establishing Transfer NR, as I have argued they cannot, then it is only dialectically fair for a critic to ask what sort of examples would help van Inwagen in this task. (For aid in this formulation of my criticism, I am indebted to the keen insight of an anonymous referee for Philosophical Review.)

23. It is worth reflecting upon the discrepancy between the Direct Argument and the Consequence Argument. At first blush, one might wonder why the argumentative strategy that I have offered here could not be applied to the Transfer of Powerlessness operator at work in the Consequence Argument (for such an attempt, see Bishop 1989, 51–59). Wouldn’t it also be true of that argument that there would be controversial entailment relations between states of an agent involved in deliberation and subsequent action? Could not one also say that these entailments presuppose that the agent in such a case could not have done otherwise? I do not think so. And this highlights
Saying Good-bye to the Direct Argument the Right Way

As I see it, Transfer NR, or van Inwagen’s Rule B, should never have been considered by the compatibilist as a legitimate inference principle in the first place. Careful thinking about the facts in virtue of which the pertinent entailments are alleged to be true, entailments such as □(P • L. ⊃ s), strongly suggests that, with applications of Rule A, these entailments lead only to what look to be innocuous inferences regarding claims of nonresponsibility, inferences that on further consideration are highly suspect. The innocuous appearance is why we compatibilists said hello to the Direct Argument in the first place. It seemed so attractive, so alluring! For it certainly seems clear, whether one is a compatibilist or an incompatibilist, that along with granting that a person is nonresponsible for the facts of the remote past and the laws of nature, one should also grant that one is nonresponsible for the fact that these result in one unique future. How, it might be thought, could a compatibilist not grant that latter point as well? It seems it can be characterized in shorthand as the claim that one is not responsible for the fact that determinism is true. But buried deep in these seemingly innocuous claims are claims of nonresponsibility that have it that one is not even partly responsible for the fact that one’s own deliberations led to one’s own actions. This, the compatibilist should not grant just because the relation is deterministic.

7. The Dialectical Fallout

What is the dialectical fallout of my rejection of Transfer NR? Consideration of van Inwagen’s positive case for Transfer NR, that is, his Rule B, and consideration of his assessment of the compatibilist’s dialectical burdens, help to point the way. Van Inwagen states that the only way he can see to argue for the inference principle is to reflect on cases like Snakebite and Plato, and to attempt to construct counterexamples to the inference principle. In his estimation, this will intuitively lead one to a dialectical asymmetry between the Direct and the Consequence Argument, one that reveals how it is that the incompatibilist is on more solid ground in advancing the Consequence Argument. Unlike the Direct Argument, the pertinent entailments at issue in the Consequence Argument are ones that, in at least one clear sense, are uncontroversially established. In relevant contexts, the agent involved in a state of deliberation could not do otherwise—contexts in which the facts and the laws are precisely fixed and it is assumed that determinism is true. (I thank Derk Pereboom for encouraging me to consider this issue and for calling to my attention Bishop’s treatment of the Consequence Argument, which indeed is structurally similar to my approach to the Direct Argument.)
see the plausibility of it since, as he sees it, there is no clear way to find any counterexamples. But, van Inwagen warns, one cannot appeal to counterexamples that assume the compatibility of moral responsibility and determinism. Framed as such, one can see how the cards are stacked against the compatibilist. If one thinks only about cases like Snakebite and Plato (when restricted to one-path cases), there transparently cannot be counterexamples to Transfer NR since the examples van Inwagen marshals all involve causal chains that never “pass through” a normally functioning agent. If, however, a compatibilist were to attempt a counterexample to a suitably restricted version of Transfer NR (like Transfer NR∗ or Transfer NR∗∗), one that did involve a causal chain passing through an agent, van Inwagen then would object that the compatibilist would be begging the question. Why? Any such case would clearly have it that a person can be responsible for what he or she does even if she or he is not responsible for what led to what was done—a patently compatibilist thesis.

Notice that van Inwagen’s manner of advancing Transfer NR turns on the presumption that the compatibilist must say something to defeat the principle. But van Inwagen himself only has the examples like Snakebite and Plato to establish Transfer NR. So why assume that the principle is sufficiently well established that a compatibilist must defeat it? My thesis is that Transfer NR, if it is to be established as a considerable threat to compatibilism in the first place, needs appropriate examples to confirm it as a compelling inference principle. Cases like Snakebite and Plato do not reveal that the inference principle can be applied to cases of agency as van Inwagen needs them to be. And if I am correct, when Transfer NR is applied to cases of normally functioning agency, one commits to highly controversial claims of nonresponsibility.

Hence, for van Inwagen to establish that Transfer NR is something that a compatibilist needs to address, it is his burden to establish the initial presumption that Transfer NR has considerable compatibilist-threatening credibility in the first place. In light of my objection, it is especially telling to revisit van Inwagen’s own characterization of the compatibilist’s dialectical burden, as I quoted it above. Recall, he (1983, 188) writes:

If the compatibilist . . . wishes to refute the argument . . . here is what he will have to do: he will have to produce some set of propositions intuitively more plausible than the validity of Rule (B) and show how these propositions entail the compatibility of moral responsibility and determinism, or else he will have to devise a counter-example to rule (B), a
counter-example that can be evaluated independently of the question whether moral responsibility and determinism are compatible.

In my estimation, there is a third option, and that is for the compatibilist to defeat the presumption that van Inwagen’s Rule B (Transfer NR) is sufficiently well established to threaten compatibilism. Rather than accepting van Inwagen’s challenge to give a nonquestion-begging counterexample to it, as Widerker, and Fischer and Ravizza do, the compatibilist can demand that van Inwagen give confirming instances of Transfer NR that establish that it applies to instances of agency. If van Inwagen cannot do that, and I have argued that he hasn’t, then it is not the compatibilist’s task to provide a counterexample to a principle that has yet to be proven a clear threat to compatibilism.

The present dialectical situation is elusive, and it is worth pausing to reflect upon it. The upshot of the preceding discussion should not be overstated. It is not that the incompatibilist concern expressed in the Direct Argument has no responsibility-challenging force. Certainly it does. There is some intuitive incompatibilist-friendly punch to the suggestion that one is not responsible for one’s actions if they are determined by factors beyond one’s control. But dialectically, Transfer NR as forecast by van Inwagen and others (including myself in earlier work) is alleged to be an inference principle that develops in a way that advances this suggestion further. Its weight is supposed to reside in ordinary judgments, ones that, we are told, reveal that basic patterns of inference about responsibility (or lack of it) have resources the incompatibilist can exploit. Doing so is supposed to allow the incompatibilist to add to his or her fundamental claim that one is not responsible for one’s actions if they are determined by factors beyond one’s control. It is this presumption that is in question. My claim is that this presumption is false. The point here is that the intuitive basis for Transfer NR, as found in cases like Snakebite and Plato, is insufficient to earn Transfer NR the status of contributing something to the Direct Argument over and above the basic incompatibilist thought that the argument means prove. Once attention is rightly drawn to causal sequences that begin with the onset of deliberation and end in action, it is far from clear that Transfer NR is a defensible inference principle.

24. For help on this point, I have profited from the astute remarks of an anonymous referee for *Philosophical Review.*
If the Direct Argument does not have the dialectical clout that on first appearance it seems to have, is it dead? No. Consider the claims of causal sufficiency that I argued created problems for the crucial nonresponsibility material conditionals, like the one that moves from Bush’s prior states of deliberation to his (allegedly) free decision to get down. It was here that I claimed that the compatibilist should point out that such claims of nonresponsibility are too controversial to be used in the service of establishing the inference principle. But they would not be if a case could be made that such causally connected states of an agent are responsibility undermining. There is a different sort of argument that incompatibilists can enlist to get this result. They can argue that if a person is manipulated into a state of deliberation through some means such as hypnosis, direct electronic manipulation of the brain, or the fancy workings of a demon, and if that state causally produces an (otherwise) free action, then the agent is not responsible for the causal relation between the deliberation and the action, and as a result, the action itself. Then the incompatibilist can argue that determinism does not differ in any significant respect from (relevant) cases of manipulation. This kind of argument, a Manipulation Argument, could be used in the service of advancing the Direct Argument, and in particular, in the service of arguing for the controversial claims of nonresponsibility that I have focused upon in my rejection of the Direct Argument.

The Manipulation Argument has figured centrally in the important incompatibilist work of Robert Kane, Derk Pereboom, and Richard Taylor, among others (Kane 1996; Pereboom 1995, 2001; and Taylor 1974). In my estimation, the Direct Argument, if it has dialectical clout, has to be bolstered by the Manipulation Argument. It is telling that, in his case for incompatibilism, in Living without Free Will, Derk Pereboom (2001, 3) appeals to a principle that is similar to the core idea in Transfer NR:

An action is free in the sense required for moral responsibility only if it is not produced by a deterministic process that traces back to factors beyond the agent’s control.

How does Pereboom establish this principle? He appeals to the Manipulation Argument. As I see it, following Pereboom’s lead, the best case for the source incompatibilist is to be made by working from the Manipulation Argument to some principle like that found in the Direct Argument. Using the Manipulation Argument, the source incompatibilist could attempt to show how in one case a manipulated agent is no
different from a determined agent. It would also have to be established that the agent is rendered nonresponsible by virtue of the manipulation. Then the source incompatibilist could appeal to a principle like Pereboom’s (or instead to Transfer NR) to generalize the results of the case study in manipulation. That, I believe, is a tough case for the compatibilist. In my estimation, it is the future of the debate, at least between source theorists, who might disagree about the compatibility of determinism and either free will or moral responsibility, but who agree that the proper focus for theorizing about free will and responsibility concerns the sources of agency and not the freedom to do otherwise.

8. Conclusion

In closing, it is worth asking, how is it that the Direct Argument acquired its prominent place in the dialectic in the first place since, as I have argued, it has yet to earn its right to be there. In earlier work, I expressed admiration for the Direct Argument. I (2001, 48) wrote, “But the basic pattern of inference offered by the Transfer NR argument for incompatibilism builds upon pellucid modal inferences. It is difficult to imagine how they could be undermined.” Indeed, it is difficult to imagine how they could be undermined if, as van Inwagen invited us to consider it, we evaluate Transfer NR by thinking about cases like Snakebite and Plato, and if we think it legitimate to restrict it to one-path cases. But in considering cases like Bush Deliberates, it quickly becomes clear that this inference principle is not at all pellucid. It is rather highly controversial and can seem otherwise only when our attention is averted from applying it directly to cases of normally functioning agents. Once we think carefully about applying Transfer NR in these cases, it is not at all clear that it is a pristine modal principle. Rather, it is an expression of an incompatibilist intuition. But then, we already knew all along that incompatibilists have thought that if our future is “written in the stars,”

25. In fact, to make perspicuous just how controversial it is, consider an analogue inference principle that would have it that claims of nonagency could be transferred. It is certainly a nonstarter to assert that if determinism is true, the fact that no agents existed at earlier moments in time could transfer through to a conclusion that no agents exist now. If determinism is true, there is a causal history according to which agents emerged from causal conditions that themselves did not include any agents. Similarly, a compatibilist holds that if determinism is true, there is a causal history according to which morally responsible agents emerged from causal conditions that themselves did not include any morally responsible agents. I am thankful to Joshua Gert for help in formulating this point.
then we cannot be accountable for it. Of course, we compatibilists have an intuitive retort to that: our future is still, at least sometimes, our making. We’ll happily settle for that as a basis for our responsibility.

References


Saying Good-bye to the Direct Argument the Right Way


