The Direct Argument Is a *Prima Facie* Threat to Compatibilism

Ori Beck[*]

Introduction

Peter van Inwagen (1983, pp. 182-188) presented a landmark argument for the incompatibility of determinism and moral responsibility. The argument—which came to be known as the *Direct Argument*—can be expressed thus:[1]

(1) Determinism is true. Hypothesis

(2) $\Box$[(LAWS & PAST) $\supset$ ACT] From (1)

(3) $\Box$[LAWS $\supset$ (PAST $\supset$ ACT)] From (2)

(4) NR[(LAWS $\supset$ (PAST $\supset$ ACT)] From (3) by rule A

(5) NR(LAWS) Premise

(6) NR(PAST $\supset$ ACT) From (4), (5) by Transfer NR[2]

(7) NR(PAST) Premise

(8) NR(ACT) From (6), (7), by Transfer NR

(9) If determinism is true, then NR(ACT) From (1)-(8)

Here, "NR(p)" abbreviates "p, and no one is, or ever has been even partly morally responsible for the fact that p". "LAWS" stands for the conjunction of all the laws of nature. "PAST" stands for some true proposition describing the complete state of the world at some time t before the existence of humans. "ACT" stands for any true proposition describing any action performed by a human agent. Lastly, "$\Box$" expresses broadly logical necessity, and "$\supset$" is the material conditional. The two rules of inference used are:

(A) $\Box p \vdash$ NR(p)

(Transfer NR) NR(p), NR($p \supset q$) $\vdash$ NR(q)

[*] I am deeply indebted to David Widerker both for encouraging me to write this paper, and for his insightful suggestions and advice. I am also heavily indebted to several anonymous reviewers for questions, criticisms and suggestions that greatly improved the paper. Further thanks go to the participants of the 2019 Free Will and Methodology conference at KU Leuven, and to the participants of the 2020 Conference of the Israeli Philosophy Association. Their comments and questions were invaluable to me. Finally, my writing was aided by the generous support of the Department of Philosophy and Kreitman School of Advanced Studies at Ben-Gurion University of the Negev.

[1] The argument is called the "Direct Argument" because it makes no explicit assumptions about the necessary conditions of moral responsibility. It thereby differs from more traditional arguments, which often make the widely doubted assumption that the possibility of alternatives is necessary for moral responsibility [for the *locus classicus* of these doubts, see Frankfurt (1969)]. The Direct Argument's name is due to Fischer and Ravizza (1998, chp. 6).

[2] The rule's name follows Fischer and Ravizza (1998, chp. 6) and Widerker (2002).

Is the Direct Argument sound? Well, premises (5) and (7), like the analytic inferences in steps (2), (3) and (9), seem difficult to dispute. It is easier to debate the uses of Transfer NR in steps (6) and (8), and the use of rule A in step (4). Here I focus on Transfer NR.[3]

The literature on Transfer NR has two main strands: one concerning Transfer NR's validity, another concerning the dialectic appropriateness of Transfer NR's use. Discussions belonging to the former strand broadly accept that the Direct Argument—and specifically, the direct argument's use of Transfer NR—poses a *prima facie* threat to compatibilism. The discussions undertake to debate Transfer NR's validity, to produce counterexamples to Transfer NR, to consider incompatibilist-friendly revisions of it, and to propose compatibilist-friendly surrogates for it.[4] The other strand of the literature on Transfer NR concerns the dialectic appropriateness of Transfer NR's use. Though this topic is already salient in Fischer's (2004), it is most straightforwardly discussed in McKenna's (2008). McKenna argues there that van Inwagen tried to establish the *prima facie* validity of Transfer NR—*in its original version*—in a dialectically improper way. He concludes that that the Direct Argument does not pose even a *prima-facie* threat to compatibilism, and that the entire former strand of the literature should thus have never developed. Or in McKenna's words:

> 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. (2008, p. 370, also see p. 377)

McKenna's arguments have not persuaded everyone. Shabo (2010b), for instance, accepts McKenna's stance with respect to certain versions of Transfer NR, but not with respect to others. More radically, Schnall and Widerker (2012) reject McKenna's position in its entirety. In a major recent response, however, Loewenstein (2016) contends that McKenna is right, and that the Direct Argument indeed does *not* pose a *prima-facie* threat to compatibilism.

My goal in this paper is modest. I will not argue that the Direct Argument is sound. Nor will I argue that Transfer NR (in any of its versions) is valid. But I will argue—against Fischer, McKenna and Loewenstein, and with Schnall and Widerker—that the Direct Argument (and specifically, the direct argument's use of Transfer NR) poses a *prima-facie* threat to compatibilism. So as I see things, it is the *second* strand of the literature which does not help us resolve the debate. To move forward, we should focus our attention on the first strand.

1 McKenna's Position

It is easiest to begin with McKenna's (2008) position. He notes that van Inwagen (1983, p. 187) attempts to establish Transfer NR's *prima facie* validity by appealing to two of Transfer NR's instances. Following McKenna (2008), I refer to the first instance as *Snakebite*:

(10) NR(John was bitten by a cobra on his 30th birthday)

(11) NR(John was bitten by a cobra on his 30th birthday $\supset$ John died on his 30th birthday)

/∴ **(12)** NR(John died on his 30th birthday)

and to the second instance as *Plato*:

[3] Objections to the validity of rule A have been made by Cyr (forthcoming), Hermes (2014a), Kearns (2011). For replies, see Robinson (2016), Turner and Capes (2018).

[4] See, e.g., Ravizza (1994), Fischer and Ravizza (1998, chp. 6), Warfield (1996), Stump and Fischer (2000), Stump (2000), McKenna (2001), Stump (2002), Widerker (2002), Fischer (2004, reprinted in his 2006), Haji (2008), Haji (2010), Shabo (2010a)Hermes (2014b), and Capes (2016). For a helpful review, see Widerker and Schnall (2014).

(13) NR(Plato died in antiquity)

(14) NR(Plato died in antiquity ⊃ Plato never met Hume)

/∴ **(15)** NR(Plato never met Hume)

Snakebite and Plato are both intuitively valid instances of Transfer NR. Therefore—van Inwagen reasons—both Snakebite and Plato are *confirming* instances of Transfer NR. They support Transfer NR's *prima facie* validity.

McKenna disagrees. He argues (2008, p. 376) that in Snakebite and Plato,

> ... 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.

According to McKenna, Snakebite and Plato are arguments that do not involve normal agency [i.e., a normally functioning agent exercising unimpaired deliberative capacities in the production of an (allegedly) free action]. Therefore, they at most support the validity of instances of Transfer NR that do not involve normal agency. But they definitely do not support the (unrestricted) validity of Transfer NR. To support Transfer NR's validity with instances, one would have to produce instances of Transfer NR that are both intuitively valid and involve normal agency. Until Transfer NR is supported in this way, compatibilists should not feel as if they have the burden of disputing Transfer NR's validity.

2 Response (A1): A Confirming Instance Involving Normal Agency

Schnall and Widerker (2012) offer McKenna two responses. First, rising to McKenna's challenge, they attempt to construct two new confirming instances of Transfer NR that involve normal agency. Second, they argue that Snakebite and Plato are perfectly legitimate ways of establishing that Transfer NR is *prima facie* valid. This section, along with section 4, presents Schnall and Widerker's new normal-agency-involving instances of Transfer NR. Their other response is discussed in section 6.

Schnall and Widerker first new instance of Transfer NR derives from the following scenario:

> One morning, the commanding officer in the army's anti-missile defence system receives a call from one of his subordinates, Jones. Jones says he is ill and cannot come in today. Unbeknown [*sic*] either to the officer or to Jones, there happens to be a man, Smith, in the area who looks exactly like Jones. Later in the day, the officer is on his way to another base when he sees Smith, whom he takes to be Jones, frolicking on the beach. The officer, after some deliberation, resolves to give Jones a severe reprimand the next day, which he does. Jones is extremely insulted by the reprimand. He denies having been on the beach, but to no avail; the officer insists, 'I saw you there with my own eyes.' (Schnall and Widerker, 2012, pp. 29-30)

Schnall and Widerker associate this scenario with the argument *Reprimand*:[5]

(16) NR(the officer believes that Jones was frolicking on the beach)

[5] Schnall and Widerker (2012) use "Reprimand" and certain other names to refer to scenarios associated with certain instances of Transfer NR. In this paper, however, I use those same names to refer to the instances of Transfer NR themselves.

(17) NR(the officer believes that Jones was frolicking on the beach $\supset$ the officer reprimands Jones)

/∴ **(18)** NR(the officer reprimands Jones)

Reprimand is clearly an instance of Transfer NR. Reprimand also involves normal agency, because the officer is a normally functioning agent exercising unimpaired deliberative capacities in the production of an (allegedly) free action. So if Reprimand is furthermore intuitively valid, it counts as a confirming instance of Transfer NR that involves normal agency - just as McKenna requires.

To establish that Reprimand is indeed intuitively valid, Schnall and Widerker (2012, p. 30) use what they dub "the reverse argument":

> Suppose someone, Alice, tells us that she thinks that the officer is blameworthy for reprimanding Jones. We might ask Alice whether she thinks that the officer should not have jumped to the conclusion that the man he saw on the beach was Jones, and therefore he is blameworthy for believing that it was Jones on the beach. Suppose she says 'No'; after all, the man he saw on the beach looked exactly like Jones. We might then ask whether she thinks that the officer is blameworthy for letting this belief lead him to reprimand Jones; perhaps he should have, instead, decided to speak to Jones in a more sympathetic way. Suppose she again answers 'No'; after all, if he believes that Jones was frolicking on the beach when he should have been defending his country, and that he must have been lying when he called in sick, then it is perfectly reasonable for the officer to conclude that Jones should be reprimanded. At this point, we will be very puzzled by Alice's position; for it seems that in this scenario, if the officer is blameworthy for reprimanding Jones... that can only be either because he is blameworthy for believing that Jones was on the beach ... or because he is blameworthy for that belief's leading him to issue the reprimand... That is, if someone denies the conclusion of this Transfer NR argument, it would be natural to assume that that person denies one of the premises – a strong indication of validity.

The reverse argument points out that we would be deeply puzzled by anyone who denied (18) but accepted both (16) and (17). Our disposition to be so puzzled indicates that Reprimand is intuitively valid.

3 A Rejoinder to (A1)?

Coming to McKenna's defense, Loewenstein (2016) argues that Reprimand does not appropriately support Transfer NR's validity. Central to her argument is what she calls "the epistemic condition on moral responsibility". According to Loewenstein (2016, pp. 215-216), the condition states that, for any morally wrong act φ, "to be morally responsible for φ-ing in a context C, an agent must understand—or at least be in a position such that she can reasonably be expected to understand—that φ-ing is morally wrong in C".

Loewenstein observes that the epistemic condition is relevant to Reprimand. This is because the officer mentioned in Reprimand neither understands, nor can be reasonably expected to understand, that it is wrong to reprimand Jones. And if the officer neither understands nor can be reasonably expected to understand that it is wrong to reprimand Jones, then by the epistemic condition the officer is not responsible for reprimanding Jones. It thus emerges that the epistemic condition explains why the officer is not responsible for reprimanding Jones. With this observation in hand, Loewenstein (2016, p. 216) writes:

> Since whether it is in principle possible for an agent to meet the *epistemic requirement* of [moral responsibility] on the assumption of determinism is not what is at issue in the dispute ... the compatibilist can appeal to a more limited principle regarding the transfer of nonculpable ignorance—that is, a principle

> which is seemingly *neutral* between compatibilism and incompatibilism—to explain why [the officer] is not responsible. And if there is an *uncontroversial* way to account for the transfer of nonresponsibility in Reprimand, then Reprimand cannot by itself justify the postulation of a further, controversial principle about [moral responsibility]. To put the point a bit differently, the compatibilist can agree with the proponent of [the Direct Argument] that a failure to meet the epistemic condition transfers through the entailment from p and $p \supset q$ to q. She can agree with this while still maintaining that, contra [Transfer NR], it is possible for an agent to be morally responsible for q even if no one meets the *control* condition to be responsible for either p or $p \supset q$.[6]

On a first reading, Loewenstein's argument may appear to be this: In Reprimand, an uncontroversial explanation of the officer's lack of responsibility for reprimanding Jones appeals to the officer's failing to meet the epistemic condition on moral responsibility. The explanation does not appeal to Transfer NR. Therefore, Reprimand does not appropriately support Transfer NR's validity.

This argument, however, is a non-sequitur. For suppose that the officer's lack of responsibility for reprimanding Jones indeed has nothing to do with Transfer NR and only to do with the officer's failure to meet the epistemic condition. In that case, we would have a Transfer NR-free explanation of the *truth of Reprimand's conclusion*, i.e., of claim (18). But it simply does not follow from this that we would have a Transfer NR-free explanation of *Reprimand's intuitive validity*, i.e., of how (18) follows from (16) and (17). And it is Reprimand's intuitive validity—not the truth of Reprimand's conclusion—that was supposed to support Transfer NR's validity. Therefore, for all that has been said so far, Reprimand's intuitive validity supports Transfer NR's validity in a dialectically proper way.

The central criticism I am making here is that it is one thing to explain the truth of an argument's conclusion, and an entirely different thing to determine whether an argument supports the validity of a certain rule of inference in a dialectically proper way. To see this more clearly, consider the argument *Sky*:

(19) The sky is blue $\wedge$ grass is green

/∴ **(20)** The sky is blue

The explanation of why the sky is blue involves certain light-scattering phenomena. It does not involve the logical properties of conjunction. Therefore, it is possible to explain why Sky's conclusion is true without appealing to (the rule of inference) Conjunction Elimination. Still, Sky's intuitive validity clearly elicits the logical intuition that Conjunction Elimination is valid. So the fact that the truth of Sky's conclusion has a Conjunction Elimination-free explanation is perfectly compatible with Sky's appropriately supporting Conjunction Elimination.

The situation with respect to Reprimand is similar: The explanation of why the officer is not responsible for reprimanding Jones may be that the officer fails to meet the epistemic condition on moral responsibility. So it may be possible to explain why Reprimand's conclusion is true without appealing to Transfer NR. Still, Reprimand's intuitive validity does elicit the logical and conceptual intuition that Transfer NR is valid. So the possibility of giving a Transfer NR-free explanation of the truth of Reprimand's conclusion is compatible with Reprimand's appropriately supporting Transfer NR's validity.

An alternative, and more charitable, way of reading Loewenstein emerges when we attend to her suggestion that there is "an uncontroversial way to account for the *transfer* of nonresponsibility in Reprimand", which appeals to a "principle regarding the *transfer* of nonculpable ignorance" stating that "failure to meet the epistemic condition *transfers through the entailment* from p and $p \supset q$ to q" (emphases added). These quotes suggest that Loewenstein's true argument is this: All parties to the discussion accept an "uncontroversial" principle regarding the transfer of

[6] By the "control condition", Loewenstein (2016, p. 215) means the principle that "to be morally responsible it is necessary that the agent have the appropriate kind of freedom, or control, over the act for which she is responsible".

nonculpable ignorance, which states that "failure to meet the epistemic condition transfers through the entailment from p and $p \supset q$ to q". This principle has two virtues. First, it explains why Reprimand is intuitively valid. Second, accepting it does not commit one to the validity of the controversial rule Transfer NR. Therefore, Reprimand's intuitive validity is explained by an uncontroversial principle that does not commit one to Transfer NR's validity. For this reason, Reprimand's intuitive validity in no way suggests that Transfer NR is itself valid.

Stated formally, the transfer principle Loewenstein must have in mind is

(Transfer NCI) $\mathrm{NCI}_{\mathrm{S}}(p), \mathrm{NCI}_{\mathrm{S}}(p \supset q) \vdash \mathrm{NCI}_{\mathrm{S}}(q)$,

where "$\mathrm{NCI}_{\mathrm{S}}(p)$" abbreviates "$p$, and S neither understands—nor can be reasonably expected to understand—that its being the case that p is morally wrong". Loewenstein's claim is that since Transfer NCI explains Reprimand's intuitive validity, Reprimand's intuitive validity does not support Transfer NR.

How does Transfer NCI explain Reprimand's intuitive validity? Presumably, Loewenstein's view is that Transfer NCI is itself intuitively valid, and therefore the following instance of it is also intuitively valid:

(21) $\mathrm{NCI}_{\text{the officer}}$(the officer believes that Jones was frolicking on the beach)

(22) $\mathrm{NCI}_{\text{the officer}}$(the officer believes that Jones was frolicking on the beach $\supset$ the officer reprimands Jones)

/∴ **(23)** $\mathrm{NCI}_{\text{the officer}}$(the officer reprimands Jones).

We take the intuitively valid inference contained in (21)-(23) to suggest that Reprimand is valid too.

However, Transfer NCI cannot explain Reprimand's intuitive validity. This is because Transfer NCI is not intuitively valid. Here is a scenario showing that it is not:

> Bob hates Carol and decides to murder her. Bob knows he can murder Carol using a Death Machine, which causes the death of the person sitting inside it. So Bob buys the machine, puts Carol inside, and flips the switch. Carol soon dies. Bob understands that Carol's death is morally wrong, but he does not care. Two further facts are important: First, the Death Machine causes Carol's death by exposing Carol to gamma radiation. Second, Bob lacks all concepts for physical phenomena such as energy, radiation, gamma rays, etc. He is thus unable to think thoughts which represent such phenomena.

The instance of Transfer NCI I associate with this scenario is *Death Machine:*

(24) $\mathrm{NCI}_{\mathrm{Bob}}$(the Death Machine exposes Carol to gamma radiation)

(25) $\mathrm{NCI}_{\mathrm{Bob}}$(the Death Machine exposes Carol to gamma radiation $\supset$ Carol dies)

/∴ **(26)** $\mathrm{NCI}_{\mathrm{Bob}}$(Carol dies)

We can see that Death Machine is invalid. To begin with, premise (24) is true: Bob is unable to think thoughts which represent physical phenomena such as energy, radiation, gamma radiation, etc. Therefore, Bob is unable to think thoughts which represent that the Death Machine exposes Carol to gamma radiation. Hence, Bob neither understands—nor can be reasonably expected to understand—that the Death Machine exposes Carol to gamma radiation. *A fortiori*, Bob neither understands—nor can be reasonably expected to understand—that it is morally wrong that the Death Machine exposes Carol to gamma radiation. Premise (25) is true for similar reasons. Since Bob is unable to think thoughts which represent physical phenomena such as energy, radiation, gamma radiation, etc., Bob is unable to think thoughts which represent that (the Death Machine exposes Carol to gamma radiation $\supset$ Carol dies). Therefore, Bob neither understands—nor can be reasonably expected to understand—that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies).[7] *A fortiori*, Bob neither understands—nor can be reasonably expected to understand—that it is morally wrong that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies). Finally, (26) is false. By stipulation, Bob understands that Carol's death is morally wrong. And since (24) and (25) are true while (26) is false, Death Machine is invalid.

Someone might try to resist this. In particular, someone might argue that if (26) is false, then (25) or something like it is false as well. After all, Bob might not understand how the Death Machine works, just as an assassin might not know exactly how the internal mechanisms of a sniper rifle work. Nevertheless, Bob does understand that (the Death Machine is activated ⊃ Carol dies). It seems that if Bob didn't understand even this much, we would have to draw the absurd conclusion that Bob is not responsible for Carol's death.

I agree with the objector that Bob understands that (the Death Machine is activated ⊃ Carol dies), and therefore that it is false that

(25*) $\mathrm{NCI}_{\mathrm{Bob}}$(the Death Machine is activated ⊃ Carol dies).

Furthermore, I am willing to agree that (25*)'s falsity is part of the explanation of (26)'s falsity. But nothing in these agreements undermines my case that Death Machine is invalid. Understanding that (the Death Machine exposes Carol to gamma radiation) and that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies) requires thinking thoughts representing gamma radiation, which Bob cannot do. Understanding that (the Death Machine is activated ⊃ Carol dies) and that (Carol dies) does not require thinking thoughts representing gamma radiation. Therefore, (24) and (25)'s truth is perfectly compatible with what the objector and I agree to, i.e., that (25*) and (26) are false. My arguments that (24) and (25) are true, while (26) is false, are therefore left unscathed. Now, since (24) and (25) are true, while (26) is false, Death Machine is invalid.

To recap: Death Machine's clear invalidity shows that Transfer NCI is not intuitively valid. Therefore, Transfer NCI cannot explain Reprimand's intuitive validity. The attempt to show that Reprimand's intuitive validity does not support Transfer NR's validity thus fails. Furthermore, since the only explanation we currently have for Reprimand's intuitive validity appeals to Transfer NR, it is perfectly reasonable to take Reprimand's intuitive validity to support Transfer NR's validity.[8]

4 Response (A2): Another Confirming Instance Involving Normal Agency

Apart from Reprimand, Schnall and Widerker seek to support Transfer NR's validity with another instance involving normal agency. They introduce this instance with the following scenario (2012, p. 31):

> Jones sees a tornado approaching. He deliberates for a few seconds as to what to do, and then decides, in order to save his life, to get into his car and drive away, and he proceeds to do so.

The instance of Transfer NR that Schnall and Widerker associate with this scenario is *Tornado*:

(27) NR(Jones sees a tornado approaching)

[7] Bob may be able to understand that every substitution of "*p*" in the context "(*p*⊃ Bob reprimands Carol)" is true. But that does not matter. Being unable to think thoughts which represent physical phenomena such as energy, radiation, gamma radiation, etc., Bob is unable to think thoughts which represent those states of affairs which the relevant substitutions of *p* would pick out. In particular, Bob is unable to think thoughts which represent that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies). Not being able to think such thoughts, Bob neither understands—nor can be reasonably expected to understand—that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies).

[8] Could one revise Transfer NCI to avoid this criticism? One suggestion in that direction would be to use Transfer NCI*: $\mathrm{NCI}^*_{\mathrm{S}}(p)$, $\mathrm{NCI}^*_{\mathrm{S}}(p \supset q) \vdash \mathrm{NCI}^*_{\mathrm{S}}(q)$, where "$\mathrm{NCI}^*_{\mathrm{S}}(p)$" abbreviates "*p*, and S neither understands—nor can be reasonably expected to understand—that *p*". Unfortunately, Death Machine shows Transfer NCI* is not intuitively valid. Bob neither understands—nor can be reasonably expected to understand—that the Death Machine exposes Carol to gamma radiation. Bob also neither understands—nor can be reasonably expected to understand—that (the Death Machine exposes Carol to gamma radiation ⊃ Carol dies). However, Bob does understand that Carol dies.

(28) NR(Jones sees a tornado approaching $\supset$ Jones drives away to save his life)

/$\therefore$ **(29)** NR(Jones drives away to save his life)

Schnall and Widerker now reason as they did with Reprimand: First, since Jones is a normally functioning agent exercising unimpaired deliberative capacities in the production of an (allegedly) free action, Tornado involves normal agency. Second, since we would be deeply puzzled by anyone who denied (29) but accepted both (27) and (28), Tornado is intuitively valid (the reverse argument). Therefore, Tornado is a confirming instance of Transfer NR involving normal agency. It meets McKenna's challenge.

5 A Rejoinder to (A2)?

Loewenstein denies that Tornado supports Transfer NR's validity in a dialectically proper way. Her argument (2016, pp. 217-218) takes the form of a dilemma: Either (i) Jones is a normally functioning agent who exercises an unimpaired deliberative capacity when he sees the approaching tornado, or (ii) not. If (i) is the case, and Jones exercises an unimpaired deliberative capacity when he sees the approaching tornado, then (at least by compatibilist lights) Jones is morally responsible for driving away to save his life. It therefore begs the question against the compatibilist to assume that Jones is not morally responsible for this. Alternatively, if (ii) is the case, and Jones' agency is undermined when he sees the approaching tornado, then Tornado does not involve *normal* agency. So no matter which horn of the dilemma one takes, it follows that Tornado does not support Transfer NR's validity in a dialectically proper way.

The dilemma, however, is unsuccessful. The fault lies in its first horn, i.e., in (i). If (i) is true, and Jones is a normally functioning agent who exercises an unimpaired deliberative capacity when he sees the approaching tornado, then—just as Loewenstein suggests—Jones is morally responsible for driving away to save his life (by compatibilist lights). It would beg the question against the compatibilist to assume otherwise. Therefore, on the assumption of (i), we should grant that (29) is false.[9] But even with this granted, it does not follow that Tornado does not support Transfer NR's validity in a dialectically proper way. To support Transfer NR in a dialectically proper way, Tornado simply needs to be an intuitively valid instance of Transfer NR that involves normal agency. That Tornado has a false conclusion [on the assumption of (i)] is perfectly compatible with that. So the dilemma argument is a non-sequitur.

Of course, it might be objected that on the assumption of (i), Tornado not only has a false conclusion, but it is also not intuitively valid. This objection has little to stand on, however. To begin, observe that on the assumption of (i), not only is Tornado's conclusion (29) false, but its premise (28) is false as well (at least by the compatibilist's lights). For if (i) is true, and Jones is a normally functioning agent who exercises an unimpaired deliberative capacity when he sees the approaching tornado, then Jones is responsible for the deliberative process that links his seeing the tornado [the antecedent of the conditional embedded in (28)] with his driving away to save his life [the consequent of the conditional embedded in (28)]. This makes it intuitive that Jones is morally responsible for the fact that (Jones sees a tornado approaching $\supset$ Jones drives away to save his life). Thus, (28) false. Another reason to hold that (28) is false on the assumption of (i), is this: Consider the disjunction ($\neg$ Jones sees a tornado approaching $\vee$ Jones drives away to save his life). The first disjunct of this disjunction is false, and the second disjunct is true. So the disjunction's only truthmaker is Jones' driving away to save his life. Now we already accepted, with Loewenstein, that on the assumption of (i), Jones is responsible for Jones' driving away to save his life. So on the assumption of (i), Jones is

[9] I doubt Schnall and Widerker would agree that (i)'s truth entails that Jones is morally responsible for fleeing the tornado. They write (2012, fn. 1) that by "*S* is morally responsible for φ-ing", they only mean that *S* is either morally blameworthy or morally praiseworthy for φ-ing. Now consider Jones' fleeing upon seeing the approaching tornado. Since that action is morally neutral, Jones is neither morally blameworthy or morally praiseworthy for performing it. So by Schnall and Widerker's lights, Jones should not be regarded as morally responsible for fleeing the tornado, even if he is a normally functioning agent who exercises an unimpaired deliberative capacity when he sees the tornado. Loewenstein (2016) responds to this objection in footnote 23 of her paper. As a debate over this response will take us too far afield, I will not pursue it further.

responsible for the disjunction's only truthmaker. It follows that on the assumption of (i), Jones is responsible for the disjunction itself. Next note that the disjunction is logically equivalent to the conditional (Jones sees a tornado approaching ⊃ Jones drives away to save his life). Assuming responsibility is preserved under logical equivalence, it follows that Jones is morally responsible for its being the case that (Jones sees a tornado approaching ⊃ Jones drives away to save his life). So again, on the assumption of (i) and by compatibilist's lights, (28) is false.

It turns out, then, that compatibilists must accept that if Jones is a normally functioning agent who exercises an unimpaired deliberative capacity when he sees the approaching tornado, both (28) and (29) are false. Compatibilists therefore lack grounds for denying that Tornado is intuitively valid. Furthermore, Schnall and Widerker's reverse argument still suggests that Tornado is valid. Loewenstein does not challenge this argument. So if take the first horn of Loewenstein's dilemma, we (including the compatibilists among us) should hold that Tornado it is an intuitively valid instance of Transfer NR that involves normal agency. And that is all that is required to satisfy McKenna's challenge to support Transfer NR in a dialectically proper way.

A different objection might now be attempted. According to it, for Tornado to constitute a "confirming instance" of Transfer NR, its premises must be true. If the premises are not all true, then Tornado is not a confirming instance of Transfer NR, and so does not support Transfer NR. This objection is misguided. When we say that an instance of a rule of inference "confirms" the rule, what we mean is that the instance of the rule suggests that the rule is *valid*. An instance can suggest that a rule of inference is valid, however, by simply being intuitively valid itself. Therefore, if Tornado is an intuitively valid instance of Transfer NR, it is also a confirming instance of Transfer NR. There is no further requirement that Tornado also have true premises. [Of course, to satisfy McKenna's challenge, Tornado must also involve normal agency. But we already secured that in taking the dilemma's first horn.]

A third and final objection might be that Schnall and Widerker's reverse argument for Tornado's validity depends on our acceptance Tornado's premises, which we cannot do when assuming that Jones is a normally functioning agent. This objection is also unconvincing. The reverse argument does not depend on our *acceptance* of Tornado's premises. The reverse argument depends only on the observation that we would be deeply puzzled by anyone who denied (29) but accepted both (27) and (28). Now, it may be true that to make this observation, one has to *hypothetically consider* scenarios in which (27) and (28) are true, and find that (29) is true in them as well. And it may also be true that in those hypothetical scenarios, Jones is not a normally functioning agent (just as the second horn of Loewenstein's dilemma suggests). But that is of little consequence. All that matters is that once one has hypothetically considered the relevant scenarios, one would observe that there is something puzzling about accepting (27) and (28) while denying (29). This would lead one, via the unchallenged reverse argument, to the conclusion that Tornado is intuitively *valid*. And once one has reached this conclusion, one is free to continue along the first horn of Loewenstein's dilemma. In so doing, one would accept that Jones is a normally functioning agent, reject both premise (28) and conclusion (29), and still insist that Tornado is intuitively valid. McKenna's challenge asks for nothing more.

6 Response (B): Snakebite and Plato Legitimately Support Transfer NR's *Prima Facie* Validity

I now turn to Schnall and Widerker's second response to McKenna (2008). It consists of an argument that Snakebite and Plato are dialectically proper ways of supporting Transfer NR's validity.

Schnall and Widerker's argument begins by noting that we would be deeply puzzled by anyone who accepted Snakebite and Plato's premises while rejecting their conclusions. This suggests, by the reverse argument, that Snakebite and Plato are intuitively valid. Schnall and Widerker (2012, pp. 32-33) add that reflection on Snakebite and Plato's intuitive validity elicits "a certain logical and conceptual intuition" in us. Specifically, it elicits the intuition that it is impossible for any arguments of Snakebite and Plato's form to have all true premises but false conclusions. Since Transfer NR plausibly encapsulates just this intuition, reflection on the arguments' validity can persuade us that Transfer NR is itself valid. Schnall and Widerker maintain that this is the dialectically proper way by which Snakebite and Plato support Transfer NR's *prima facie* validity.

Schnall and Widerker (2012, pp. 33-34) also compare the way in which van Inwagen's tried to support Transfer NR's validity with the way in which Singer (1972) tried to support the truth of the moral principle *Prevent*:[10]

> If it is in our power to prevent something bad from happening, without thereby sacrificing anything of comparable moral importance, we ought, morally, to do it.

Singer (1972, p. 231) tried to support Prevent's truth by citing an uncontroversial confirming example. The example was that if one can save a child from drowning in a shallow pool at the mere sacrifice of getting one's clothes muddy, then one ought to do it. This example was meant to elicit in us precisely the moral intuition captured in Prevent. Later on in his paper, Singer used Prevent to establish a more controversial thesis. In particular, he used Prevent to argue that people are morally obligated to give up their earnings until the point of equal marginal utility.

Schnall and Widerker note that Singer's strategy is analogous to van Inwagen's: van Inwagen tried to support Transfer NR's validity by citing uncontroversial confirming instances, i.e., Snakebite and Plato. These examples were meant to elicit in us the "logical and conceptual intuition" captured in Transfer NR. Later on, van Inwagen used Transfer NR to establish a more controversial thesis. In particular, he used Transfer NR to establish that if determinism is true, no agent is morally responsible for acting as she did.

Having shown that Singer's and van Inwagen's argumentative strategies are closely analogous, Schnall and Widerker (2012, pp. 33-34) further note that Singer's argument is surely not dialectically improper. Rather, the argument places the burden of proof on the shoulders of Prevent's opponents. They believe the same is true (*mutatis mutandis*) of van Inwagen.

A possible objection to Schnall and Widerker's argument is that talk of "intuiting the validity" of an inference does not make sense. When we are talking about a single claim, the objection goes, it makes sense to ask whether it seems true or not. In contrast, when we are talking about an inference, it makes no sense ask whether it seems valid or not. Appealing to an inference's "seeming valid" is, in short, just unintelligible, or at least confused.[11]

My response is that I see no plausibility in the view that there can be no intuitions (i.e., intellectual seemings) to the effect that a given inference X is valid. Not only does the objection fail to provide any grounds for accepting such a view, there are also strong reasons for rejecting it: First, whatever we can be inclined to believe, we can also intuit.[12] So, since we can be inclined to believe that a given inference X is valid, it can also seem to us that inference X is valid. Second, I can just introspectively tell that certain inferences seem valid to me. For example, the inference involved in the argument Sky (section 3), and even the inference rule Conjunction Elimination itself (which Sky employs), seem valid to me. So that's proof by counterexample. Third, it had better be the case that inferences can seem valid to us. Otherwise, the only way to support an inference's validity would be by inferring that the inference is valid from some premises. But to do this second inferring in an epistemically respectable fashion, we would need some assurance of the second inference's validity. And if this assurance could only be provided by a third argument, a regression would begin. This realization leads to a question familiar from other regression arguments in epistemology. The question is this: How would the regression end? Well, two possibilities that are not even remotely appealing are that the regression would run infinitely long, or that it would end up going in a circle (in which one inference's validity ultimately supports the validity of another, which in turn supports it). A third possibility is that the regression

[10] The name is due to Loewenstein (2016, p. 220).

[11] I am grateful to an anonymous reviewer for encouraging me to discuss this objection.

[12] In fact, on some views, to intuit just *is* to be inclined to believe. For extended discussions, see Tucker's (2013) excellent collection.

terminates with an inference whose validity cannot be supported by anything. That too is not remotely appealing. The only possibility left open, then, is that the regression terminates with an inference that is supported just by its seeming valid. This possibility is perfectly fine. But it does require that inferences can seem valid to us.

If inferences can seem valid to us (as I have been arguing), there should be no reason to doubt that we can report on their so seeming by making reports like "inference X seems valid to me". Thus, the objection that it is unintelligible or confused to appeal to an inference's "seeming valid", fails.

7 A Rejoinder to (B)?

Despite the preceding discussion, Loewenstein (2016) argues that Snakebite and Plato are dialectically improper ways of supporting Transfer NR's validity.

Consider the conclusion of Snakebite (roughly, that no one is morally responsible for John's death) and the conclusion of Plato (roughly, that no one is morally responsible for Plato's never meeting Hume). Loewenstein notes that the compatibilist can explain why these conclusions are true. The explanation is that neither John's death nor Plato's never meeting Hume are caused in a way that appropriately "passes through" a normally functioning agent. Therefore, nobody is responsible for these events. Now, the important point is that this explanation does not commit one to Transfer NR's validity. Hence, the compatibilist can explain why Snakebite and Plato's conclusions are true while also rejecting Transfer NR. According to Loewenstein, it follows from this that it was dialectically improper for van Inwagen to support Transfer NR's validity with Snakebite and Plato.

To better explain this argument, Loewenstein extends Schnall and Widerker's analogy between Singer's and van Inwagen's argumentative strategies. She begins by supposing that Singer is arguing against an opponent, Brown. Brown has a rich and well-defended moral theory which centers on one single, overarching moral principle, M. Like Singer's principle Prevent, M entails that you should save children from drowning in shallow pools (under certain circumstances). Therefore, both Brown and Singer have explanations of why you should save children from drowning in shallow pools (under certain circumstances). Furthermore, Brown's and Singer's explanations are mutually inconsistent. This is because M entails that people are *not* morally obligated to give up their earnings until the point of equal marginal utility—in contradiction to Singer's principle Prevent. Evaluating this situation, Loewenstein remarks that the fact that you should save children from drowning in shallow pools (under certain circumstances) "gives no more support to Singer's principle [Prevent] than it does to Brown's [M]" (2016, p. 222). She therefore concludes that it would be dialectically improper for Singer to support Prevent's truth with the drowning case.

Loewenstein takes this dialectical situation to be analogous with van Inwagen's dialectical situation in his debate with the compatibilist: The compatibilist has an explanation of why Snakebite and Plato's conclusions are true. The explanation is that neither John's death nor Plato's never meeting Hume are caused in a way that appropriately "passes through" a normally functioning agent. This explanation does not commit one to regard Transfer NR as valid. Therefore, van Inwagen and the compatibilist have mutually inconsistent explanations of why Snakebite and Plato's conclusions are true. Evaluating this situation, Loewenstein feels that the truth of Snakebite and Plato's conclusions is explained by the compatibilist at least as well as it is explained by van Inwagen. From this she concludes that it is dialectically improper to support Transfer NR with Snakebite and Plato. In her words (2016, p. 222):

> ...it is dialectically inappropriate for van Inwagen to use Snakebite ... against opponents who have an alternative explanation for why no one is morally responsible for John's death (that is, that the causal chain terminating in John's death does not pass through a normally functioning agent in the requisite way).

But Loewenstein's reasoning is flawed. We might agree that the compatibilist can explain why Snakebite and Plato's conclusions are true while rejecting Transfer NR. But as we saw in section 3, it simply does not follow from this that it is dialectically improper to support Transfer NR with Snakebite and Plato. Recall our discussion of the argument Sky [claims (19)-(20)]. It clearly illustrates that explaining the truth of certain conclusions is one thing; and that determining whether certain arguments support the validity of a certain rule of inference in a dialectically proper way is quite another thing. This general lesson has application in the case of Snakebite and Plato. The explanation of why nobody is responsible for John's death may be that the causal chain leading to the death does not appropriately "pass through" a normally functioning agent. So it may be possible to explain why Snakebite's conclusion is true without appealing to (the rule of inference) Transfer NR. Still, Snakebite's intuitive validity does elicit the logical and conceptual intuition that Transfer NR is valid. So the possibility of giving a Transfer NR-free explanation of the truth of Snakebite's conclusion is compatible with Snakebite's appropriately supporting Transfer NR.

A related issue arises for Loewenstein's extended analogy between Singer's and van Inwagen's argumentative strategies. To see this, consider Brown's argument against Singer: The fact that one should save children from drowning in shallow pools (under certain circumstances) supports the moral principle M at least as well as it supports the moral principle Prevent; therefore, it is dialectically improper to support Prevent with the fact that one should save children from drowning in shallow pools (under certain circumstances). This argument has some weight. But what should be the compatibilist's parallel argument against van Inwagen? Here the thing to note is that while the dispute between Singer and Brown concerns the *truth* of rival moral principles, the dispute between van Inwagen and the compatibilist concerns the *validity* of rival rules of inference. Therefore, where Brown claims that the truth of M is supported by the obligation to save children from drowning in shallow pools (under certain circumstances) at least as well as Prevent's truth, the compatibilist should claim that the *validity* of some compatibilist-friendly rule of inference is supported by Snakebite and Plato's intuitive validity at least as well as Transfer NR's validity. So the parallel argument that the compatibilist should make against van Inwagen is this: Snakebite and Plato's intuitive validity supports the validity of some rival compatibilist-friendly rule of inference at least as well as it supports Transfer NR's validity; therefore it is dialectically improper to support Transfer NR's validity with Snakebite and Plato's intuitive validity.

Loewenstein does not make this last argument, however. As her above quotation shows, she focuses on the claim that the compatibilist can explain why Snakebite and Plato's conclusions are *true* while rejecting Transfer NR. But this is not pertinent. If Loewenstein wants to argue that Snakebite and Plato constitute dialectically improper ways of supporting Transfer NR's validity, then in lieu of explaining the truth of these arguments' conclusions in a Transfer NR-free way, she must provide some compatibilist-friendly rule of inference which is supported by Snakebite and Plato's intuitive validity at least as well as Transfer NR.

Perhaps anticipating this concern, in a long footnote at the end of her paper Loewenstein offers a different argument for holding that Snakebite and Plato are dialectically improper ways of supporting Transfer NR's validity. She writes (2016, fn. 28):

> I agree that a satisfying compatibilist-friendly story should account for the apparent validity of cases like Snakebite and Plato; however, it seems to me that the story on offer here does just that. ... If the demand is for an explanation for *why* the premises (at least appear to) entail that there is no appropriate agent involved in the right kind of way, consider Snakebite again. The first premise says that no one is responsible for the bite. From this it follows that the causal chain resulting in the bite does not pass through a normally functioning agent (at least not in the right kind of way for someone to be responsible for John's death in virtue of being responsible for the bite). The second premise says that no one is responsible for the fact that the bite results in John's death. From this it follows that the causal chain

> linking the bite to John's death does not pass through a normally functioning agent (at least not in the right kind of way for someone to be responsible for the fact that the bite results in John's death). Finally, upon reflection of the premises we can infer that the causal chain does not pass through an agent at the moment of the bite itself. Since snakes are not agents (at least not in the relevant sense), the snakebite is not itself an agent-involving link in the chain. ... If there is no appropriate agent involved in the causal chain leading up to the bite, and if there is no appropriate agent involved in the causal chain linking the bite to John's death, and if the bite itself is not an agent-involving event, then it seems to follow intuitively that there is no appropriate agent involved in the causal chain leading up to John's death (and thus, that no one is responsible). (Note that whether or not the inference is actually valid is irrelevant. All that is needed to meet the demand, here, is a compatibilist-friendly explanation of the intuitive *appearance* of validity.)

In this passage, Loewenstein attempts to provide a compatibilist-friendly explanation of Snakebite's intuitive *validity*. The explanation has three parts: The first part consists in the suggestion that Snakebite's premises (10) and (11)—roughly stating that nobody is responsible for the fact that John was bitten by a cobra on his 30th birthday, and that nobody is responsible for the fact that (John was bitten by a cobra on his 30th birthday ⊃ John died on his 30th birthday)—entail a certain conclusion (or a certain set of conclusions), T. The second part consists in the suggestion that conclusion T intuitively entails (roughly) that no appropriate agent is involved in the causal chain leading up to John's death. Finally, the third part roughly consists in the suggestion that, intuitively, if no appropriate agent is involved in the causal chain leading up to an event, then no agent is morally responsible for the event itself. By taking John's death to be the event in question, we can see how Snakebite's original conclusion (12)—roughly stating that nobody is responsible for the fact that John died on his 30th birthday—is meant to follow.

This significantly improves Loewenstein's position. In particular, Loewenstein's position now promises to help us identify a compatibilist-friendly rule of inference that stands as an alternative to Transfer NR. But before I evaluate this position, I must raise a point of interpretation.

In spelling out the conclusion T that putatively follows from Snakebite's premises, Loewenstein writes that the premises (*inter alia*) entail that "the causal chain resulting in the bite does not pass through a normally functioning agent (at least not in the right kind of way for someone to be responsible for John's death in virtue of being responsible for the bite)." The parenthetical remark is important here. For suppose that a strong earthquake occurs where John—a normally functioning agent—happens to be standing. The earthquake in no way compromises John's normal functioning, but it does cause John to inadvertently trip on the snake, which in turn causally contributes to the snake's later biting John. Under this supposition, it is strictly true to say that the causal chain resulting in the bite passes through a normally functioning agent. But the chain does so in an uninteresting way. Specifically, it passes through a normally functioning agent, but not in virtue of the agent's *exercising* their capacity for normal agential functioning. Loewenstein's parenthetical remark (along with other, similar, remarks) makes clear that—by Loewenstein's lights—Snakebite's premises might be consistent with the possibility that certain causal chains in the case *do* pass through a normally functioning agent in an uninteresting way (as happens in the earthquake case). All that Loewenstein insists on is that Snakebite's premises entail that the those causal chains do *not* pass through a normally functioning agent in some other, more interesting, way.

So what precisely is the way *i* such that, according to Loewenstein, Snakebite's premises entail that certain causal chains in the case do not pass through a normally functioning agent in way *i*? There are two options, corresponding to two readings of Loewenstein's parenthetical appeal to "causal chains that pass through a normally functioning agent in the right kind of way for someone to be responsible for *x*". On the first reading, the last expression picks out causal chains that pass through a normally functioning agent in virtue of the agent's exercising their capacity for normal agential functioning. So on this reading, Snakebite's premises entail that certain causal chains in the case do not pass through a normally functioning agent in virtue of their exercising their capacity for normal agential functioning. On the second reading, the expression picks out causal chains that pass through normally functioning agents who are morally responsible for *x*. Since non-normally functioning agents cannot be morally responsible for *x*, this is equivalent to saying that the expression picks out causal chains that pass through agents who are morally responsible for *x*. So on the second reading, Snakebite's premises entail that certain causal chains in the case do not pass through a normally functioning agent who is morally responsible for *x*.[13]

I do not know which of these two readings is more faithful to Loewenstein's original intention. Moreover, different philosophers I have discussed this with seem to have very strong, but also very different, beliefs on the matter (differences that do *not* track the compatibilist-incompatibilist divide). As a result, I must evaluate Loewenstein's compatibilist-friendly explanation of Snakebite's intuitive validity on both readings. My evaluation will reveal that neither reading leads to a successful compatibilist-friendly explanation of Snakebite's intuitive validity. The flaw with the first reading will be that it specifies conclusion T in such a way that the conclusion does not follow from Snakebite's premises. The flaw with the second reading will be that it either relies on false claims, or entails the falsity of compatibilism (and so is not compatibilist-friendly).

Start with the first reading, on which the explanation of Snakebite's intuitive validity is this: First, the conclusion T which is suggested to follow from Snakebite's premises (10) and (11) is the conclusion that

(30)(a) no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain leading up to the bite,

(30)(b) no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain linking the bite to John's death, and

(30)(c) the bite itself is not an agent-involving event.

Second, it is suggested that (30)(a)-(c) intuitively entail that no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain leading up to John's death. The rule of inference which would this underwrite this entailment is the rule

(Transfer NA) NA1(*p*), NA2(*p* causes *q*), NA3(*p*) $\vdash$ NA1(*q*),

where "NA1(*p*)" abbreviates "*p*, and no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain leading up to its being the case that *p*"; "NA2(*p* causes *q*)" abbreviates "its being the case that *p* causes its being the case that *q*, and no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain linking its being the case that *p* with its being the case that *q*"; and NA3(*p*) abbreviates "*p*, and its being the case that *p* is not an agent-involving event".

Third, (30)(a)-(c) entail—by Transfer NA—that no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain leading up to John's death. To complete the explanation of Snakebite's intuitive validity, we need a principle to connect this last entailment with Snakebite's conclusion (12). This principle would be:

(P) Intuitively, if no appropriate agent is, in virtue of their exercising their capacity for normal agential functioning, involved in the causal chain leading up to an event, then no agent is morally responsible for the event itself.

Together, it follows from (P) and the last entailment that no agent is morally responsible for John's death. This is (roughly) Snakebite's conclusion.

[13] I am grateful to an anonymous referee for recommending this second reading to me.

As aforementioned, this explanation of Snakebite's intuitive validity is unsuccessful. The trouble is that Snakebite's premises do not entail (30)(a)-(c). To begin, they do not entail 30(a). The premises are, for instance, consistent with the possibility that John pokes the snake, which annoys it and thus causally contributes to its later biting John. Since the poke is an exercise of John's capacity for normal agential functioning, John counts as a normally functioning agent involved in the causal chain leading up to the bite in virtue of exercising his capacity for normal agential functioning. At the same time, John is not morally responsible for the bite, because he justifiably believes that snakes cannot bite, and actually enjoy being poked. Furthermore, Snakebite's premises do not entail (30)(b). For example, the premises are consistent with the possibility that after the bite occurs, John decides not to take what is in fact anti-venom which would save his life. Since the decision is an exercise of John's capacity for normal agential functioning, John counts as a normally functioning agent involved in the causal chain linking the bite to John's death in virtue of exercising his capacity for normal agential functioning. At the same time, John is not morally responsible for the fact that (John was bitten by a cobra on his 30th birthday $\supset$ John died on his 30th birthday), because he justifiably believes, with respect to the life-saving anti-venom, that it would have no medical effect on him.[14]

Turn then to the second reading of Loewenstein's explanation of Snakebite's intuitive validity. On it, the explanation is this: First, the conclusion T which purportedly follows from Snakebite's premises (10) and (11) is the conclusion that

(31)(a) no agent who is morally responsible for the bite is involved in the causal chain leading up to the bite,

(31)(b) no agent who is morally responsible for the causal chain linking the bite to John's death is involved in that causal chain, and

(31)(c) the bite itself is not an agent-involving event.

Second, it is suggested that (31)(a)-(c) intuitively entail that no agent who is morally responsible for John's death is involved in the causal chain leading up to it. The rule of inference which would this underwrite this entailment is the rule

(Transfer NA*) NA1*(p), NA2*(p causes q), NA3*(p) $\vdash$ NA1*(q),

where "NA1*(p)" abbreviates "p, and no agent who is morally responsible for its being the case that p is involved in the causal chain leading up to its being the case that p"; "NA2*(p causes q)" abbreviates "its being the case that p causes its being the case that q, and no agent who is morally responsible for the causal chain linking its being the case that p to its being the case that q is involved in that causal chain"; and NA3*(p) abbreviates "p, and its being the case that p is not an agent-involving event".

Third, (31)(a)-(c) entail—by Transfer NA*—that no agent who is morally responsible for John's death is involved in the causal chain leading up to it. To complete the explanation of Snakebite's intuitive validity, we need a principle to connect this last entailment with Snakebite's conclusion (12). This principle would be:

(P*) Intuitively, if no agent who is morally responsible for an event is involved in the causal chain leading up to the event, then no agent is morally responsible for the event itself.

Together, it follows from (P*) and the last entailment that no agent is morally responsible for John's death. Again, this is (roughly) Snakebite's conclusion.

This explanation of Snakebite's intuitive validity is either unsuccessful or *not* compatibilist-friendly. According to compatibilists, recall, agents in deterministic worlds can be morally responsible. So let's follow them and assume

[14] Seth Shabo raised closely related points during the discussion of Loewenstein's work at the APA Pacific Division Meeting of 2013.

that *w* is a deterministic world, that *m* is the *temporally first* event in *w* for which an agent is morally responsible, and that *b* is some event (i) which occurred in *w* both before *m* and before the existence of agents, and (ii) which caused *m*. Let "M" stand for a proposition describing *m*, and let that "B" stand for a proposition describing *b*. Now consider the following instance of Transfer NA*, argument *Q*:

(32) NA1*(B)

(33) NA2*(B causes M)

(34) NA3*(B)

/∴ **(35)** NA1*(M)

Premise (32) is true at *w*. After all, *b* occurred in *w* before the existence of agents. Therefore, no agent is involved in the causal chain leading up to *b*. *A fortiori*, no agent who is morally responsible for *b* is involved in the causal chain leading up to *b*. Premise (33) is also true at *w*. Event *b* does cause *m*, and the causal chain linking *b* to *m* clearly occurred before *m*. So, since *m* is the *first* event in *w* for which an agent is morally responsible, no agent is morally responsible for the causal chain linking *b* to *m*. It follows that no agent who is morally responsible for the causal chain linking *b* to *m* is involved in that causal chain. Finally, premise (34) is true at *w*. Event *b* occurred in *w* before the existence of agents, and so *b* is not an agent-involving event.

Since argument *Q*'s premises are true at *w*, and since Transfer NA* is (purportedly) at least *intuitively* valid, conclusion (35) is itself at least intuitively true at *w*. In other words, it is intuitive that at *w*, no agent who is morally responsible for *m* is involved in the causal chain leading up to *m*. So by (P*), it is also intuitive that no agent is morally responsible for *m*. But it is *not* intuitive that at *w*, no agent is morally responsible for *m*! We have, after all, *defined m* as the temporally first event in *w* for which an agent *is* morally responsible. Contradiction.

How did we reach this contradiction? Well, we reached it by relying on the three ideas: (i) Agents in deterministic worlds can be morally responsible. (ii) Transfer NA* is intuitively valid. (iii) Principle (P*) is true. One of these three ideas must be false. If (i) is false, then compatibilism is false. If (ii) or (iii) are false, then a claim necessary for the present explanation of Snakebite's intuitive validity is false. Either way, we lack a compatibilist-friendly explanation of Snakebite's intuitive validity that rivals the explanation given by Transfer NR. Furthermore, *if* the new explanation of Snakebite's intuitive validity succeeds, it does so at the cost of replacing Transfer NR with inferential rules that are no less damaging to compatibilism than Transfer NR itself.

8 Fischer's Position

Let's take stock. This paper opened with McKenna's position, on which discussions of Transfer NR's validity "should never have even gotten off the ground" (2008, p. 370). The paper then discussed Schnall and Widerker's (2012) replies, along with Loewenstein's (2016) rejoinders. The discussion suggested that Loewenstein and McKenna have not successfully established their positions. Along the way, however, a related (even if somewhat different) dialectical point became salient. Specifically, it became salient that the compatibilist *can* argue that Transfer NR does not pose even a *prima facie* threat to compatibilism. However to do so, the compatibilist should show that the inference rule instances which have been provided to support Transfer NR's validity also support the validity of some rival, compatibilist-friendly, inference rule at least as well.

Some readers will now have an odd *deja vu* feeling. This is because the last claim closely resembles a claim that Fischer has made on the compatibilist side, even before McKenna, Schnall and Widerker, and Loewenstein argued for their present positions. I therefore turn to Fischer's view. In so doing, I also turn away from van Inwagen's original version of Transfer NR and consider more contemporary versions of the rule.

Unlike McKenna and Loewenstein, Fischer does not hold that discussions of Transfer NR's validity should not have gotten off the ground. In fact, Fischer (in Fischer & Ravizza, 1998, chp. 6) has made invaluable contributions to discussions of Transfer NR's validity, which include producing preemption- and simultaneous overdetermination-counterexamples (sometimes called "2-path-counterexamples") to Transfer NR. These counterexamples led McKenna (2001), Stump (2000, 2002) to revise Transfer NR, essentially by restricting it to non-preempted and non-overdetermined cases (sometimes called "1-path cases"). It is in his (2004) response to these new versions of Transfer NR that Fischer makes the dialectical point presently at issue.

Fischer argues that the debate between the compatibilist and the incompatibilist over the Direct Argument has reached a "dialectical stalemate". He (2004, p. 199) concedes, however, that to show that a dialectical stalemate has been reached, one must argue that the inference rule instances provided to support McKenna (2001), Stump (2000) and Stump's (2002) versions of Transfer NR also support to some rival, compatibilist-friendly, inference rule at least as well. (This is consistent with my own suggestions above.) Fischer then undertakes to construct just such a rule. This rule is (2004, pp. 201-202):

(Transfer NRC) NR(p), NR($p \supset q$), on the actual path that leads from its being the case that p to its being the case that q, 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 (e.g., a factor that distorts or impairs the distinctive process of human practical reasoning) $\vdash$ NR(q).

Fischer (2004) notes that Transfer NRC is a compatibilist-friendly rule, which cannot reasonably be used as part of a direct argument for incompatibilism. Specifically, in the Direct Argument above, Transfer NRC cannot justify the inferential step

(8) NR(ACT) From NR(PAST $\supset$ ACT) and NR(PAST)

without requiring an instance of the rule's third premise-scheme, which the compatibilist could reasonably reject.

It is an important question whether the inference rule instances provided to support McKenna (2001), Stump (2000) and Stump's (2002) versions of Transfer NR also support Transfer NRC at least as well. Still, this is not quite the question I will discuss here.[15] This is because the current state-of-the-art incompatibilist-friendly version of Transfer NR is a rule due to Capes. Capes' version says that "if a person is not even partly morally responsible for *any* of the circumstances that led to a particular outcome, and if that person is not even partly morally responsible for the fact that those circumstances led to that particular outcome, then the person is not even partly morally responsible for the outcome in question either" (2016, p. 1484). In Capes' (2016, pp. 1484-1488, 1491) formal notation:

(B*) $\mathrm{NR}_S(C_p)$, $\mathrm{NR}_S(C_p \supset p) \vdash \mathrm{NR}_S(p)$,

where "$\mathrm{NR}_S(p)$" abbreviates "p, and agent S is not even partly morally responsible for its being the case that p", and "C_p" is a proposition describing all the antecedent circumstances that led to its being the case that p, including the laws of nature. Capes' (2016, pp. 1490-1492) uses B* as part of a direct argument for incompatibilism, thereby demonstrating that B* is incompatibilist-friendly. Furthermore, we shall see that B*'s validity can be supported not just by the intuitive validity of (variants of) Snakebite and Plato, but also by the intuitive validity of (variants of) Reprimand and Tornado. So, given B*, we may set aside the older versions of Transfer NR suggested by McKenna (2001), Stump (2000, 2002).

[15] However, for an interesting discussion bearing on this question, see McKenna (2008, pp. 365-370).

If we wish to determine whether the latest incompatibilist-friendly version of Transfer NR poses a *prima facie* threat to compatibilism, the precise question we must ask is this: Do the instances that support B* also support Transfer NRC at least as well? I discuss this question next.

9 Instances Support B* Better Than They Support Transfer NRC

Two arguments can be made to show that the instances that support B* do so to a greater extent than the extent to which they support Transfer NRC. The first such argument is short and simple; the second more involved.

The short argument is this: Unlike B*, Transfer NRC explicitly invokes the concept of a "factor that *uncontroversially* undermines moral responsibility". Therefore, anyone attempting to appreciate Transfer NRC's validity must take into account the historical or sociological controversy between compatibilists and incompatibilists. But the conceptual and logical intuitions that compatibilists or incompatibilists entertain with respect to rules concerning the transfer of non-responsibility plausibly only involve the concepts of agency, moral responsibility, causality, law, etc. It is highly implausible that they involve consideration of the historical or sociological state of any controversy that is downstream of these concepts. For this reason, it is implausible that compatibilists or incompatibilists ever have any logical or conceptual intuitions regarding Transfer NRC's validity. There is no similar reason to doubt that compatibilists or incompatibilists have the logical and conceptual intuition that B* is valid. It is therefore plausible that B*-instances do elicit the intuition that B* is valid, but do not elicit the intuition that Transfer NRC is valid.

The more involved argument to the same conclusion requires us to look carefully at the particular instances that might be used to support B*. These might be intuitively valid instances of B* inspired by Snakebite, Plato, Reprimand and Tornado.[16] For brevity, I will focus only on the instance of B* inspired by Snakebite. But it should be clear from my discussion that all my points generalize to cover the other instances of B* as well.

The Snakebite-inspired instance of B * is *Snakebite**:

(36) $\mathrm{NR}_{\mathrm{John}}(C_{\text{John died on his 30th birthday}})$

(37) $\mathrm{NR}_{\mathrm{John}}(C_{\text{John died on his 30th birthday}} \supset \text{John died on his 30th birthday})$

/∴ **(38)** $\mathrm{NR}_{\mathrm{John}}(\text{John died on his 30th birthday})$

Schnall and Widerker's (2012, p. 30) reverse argument suggests that Snakebite* is intuitively valid: Suppose someone were to tell you that John is not even partly morally responsible for any of the circumstances that led to his death (e.g., he is not responsible for his having been bitten by a cobra on his 30th birthday, or for the relevant laws of nature), or for the fact that those circumstances did indeed lead to his death. If the person then insisted that John *is* nonetheless responsible for his death, you would be deeply puzzled. Your disposition to be so puzzled indicates that *Snakebite** is intuitively valid.

Snakebite*'s intuitive validity plausibly supports B*'s validity. As Schnall and Widerker (2012) would argue, by reflecting on Snakebite*'s intuitive validity it becomes intuitive to us that it is impossible for any arguments of Snakebite*'s form to have all true premises but false conclusions. Since B* encapsulates just this intuition, reflection on Snakebite* can persuade us that B* is valid.

The remaining question is whether Snakebite*'s intuitive validity supports Transfer NRC's validity at least as well as it supports B*'s validity. For this to be the case, Snakebite*'s intuitive validity would need to be due (or otherwise appropriately related) to the intuitive validity of some Snakebite*-like instance of Transfer NRC, and the intuitive validity of this Snakebite*-like Transfer NRC-instance would need to support Transfer NRC's validity to the

[16] Capes' (2016, pp. 1484-1489) also carefully supports B*.

appropriate degree. Now, the only Snakebite*-like instance of Transfer NRC to which Snakebite*'s intuitive validity could plausibly be due (or otherwise appropriately related) is *Snakebite*$^{\mathrm{NRC}}$:

(39) NR($C_{\text{John died on his 30th birthday}}$)

(40) NR($C_{\text{John died on his 30th birthday}} \supset$ John died on his 30th birthday)

(41) On the actual path that leads from any of the antecedent circumstances (including the laws of nature) that led to its being the case that John died on his 30th birthday to its being the case that John died on his 30th birthday, 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 (e.g., a factor that distorts or impairs the distinctive process of human practical reasoning)

/∴ **(12)** NR(John died on his 30th birthday)

I grant that Snakebite$^{\mathrm{NRC}}$ is intuitively valid. We would be deeply puzzled by anyone who denied (12) but accepted (39)-(41). Nevertheless, I deny Snakebite$^{\mathrm{NRC}}$ supports Transfer NRC's validity in the way that Snakebite* supports B*'s validity. My argument for this will proceed as follows: By considering a bit of reasoning closely related to Snakebite$^{\mathrm{NRC}}$, I will show that Snakebite$^{\mathrm{NRC}}$'s premise (41) is argumentative "dead weight", and that Snakebite$^{\mathrm{NRC}}$ does not elicit the intuition that Transfer NRC is valid. From this I will conclude that Snakebite$^{\mathrm{NRC}}$ does not support Transfer NRC's validity in the distinctively direct way in which Snakebite* supports B*'s validity.

The bit of Snakebite$^{\mathrm{NRC}}$-related reasoning that I wish to consider is the argument *Snakebite*$^{\mathrm{NR}\neg\mathrm{C}}$:

(39) NR($C_{\text{John died on his 30th birthday}}$)

(40) NR($C_{\text{John died on his 30th birthday}} \supset$ John died on his 30th birthday)

(¬41) ¬[On the actual path that leads from any of the antecedent circumstances (including the laws of nature) that led to its being the case that John died on his 30th birthday to its being the case that John died on his 30th birthday, 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 (e.g., a factor that distorts or impairs the distinctive process of human practical reasoning)]

/∴ **(12)** NR(John died on his 30th birthday)

Just like Snakebite$^{\mathrm{NRC}}$, Snakebite$^{\mathrm{NR}\neg\mathrm{C}}$ is intuitively valid. We would be deeply puzzled by anyone who denied (12) but accepted (39), (40) and (¬41). In accepting (39), (40) and (¬41), one would be accepting that nobody is even partly morally responsible for any of the antecedent circumstances (including the laws of nature) that led to John's death, and that nobody is even partly morally responsible for the fact those circumstances led to John's death. If anyone accepted this but nevertheless insisted that someone is at least partially morally responsible for John's death, then—notwithstanding one's further acceptance of (¬41)—we would find that person's position deeply puzzling. Therefore, by Schnall and Widerker's (2012) reverse argument, intuition is on the side of Snakebite$^{\mathrm{NR}\neg\mathrm{C}}$'s being valid.

One might initially be inclined to reply that there is some incoherence involved in accepting (39) and (40) on the one hand, and accepting (¬41) on the other hand. One might feel that this incoherence makes our intuitions about the validity of Snakebite$^{\mathrm{NR}\neg\mathrm{C}}$ unreliable. But this reply is doubly false. First, all arguments with jointly incoherent premises are (trivially) valid. So if there were some incoherence involved, the validity of Snakebite$^{\mathrm{NR}\neg\mathrm{C}}$ would be guaranteed. Second, and more importantly, *no* incoherence is involved in jointly accepting (39), (40) and (¬41). This is because (39) and (40) are compatible with the possibility that, on the one hand, some factor that could be thought to ground moral responsibility is on the path that leads to John's death, while on the other hand, factors that uncontroversially undermine agents' moral responsibility for the death exist, but are all factors that play no causal role in John's death, and therefore that are not on the path that leads to John's death.

We have seen, then, that both Snakebite$^{\text{NRC}}$and Snakebite$^{\text{NR}\neg\text{C}}$ are intuitively valid. This has a simple explanation: An argument whose *only* premises are (39) and (40) and whose conclusion is (12) is by itself intuitively valid. This can be independently verified by yet another appeal to the regress argument: We would be deeply puzzled by anyone who denied (12) but accepted (39)-(40). Now, given the monotonicity of logical inference, the intuitive validity of the inference from (39)-(40) to (12) is sufficient to ensure Snakebite$^{\text{NRC}}$and Snakebite$^{\text{NR}\neg\text{C}}$'s intuitive validity. In other words, neither accepting (41) nor accepting ($\neg$41) is necessary for (12) to intuitively follow from (39) and (40). In this sense, (41) and ($\neg$41) are *dead weight* in the validity of Snakebite$^{\text{NRC}}$and Snakebite$^{\text{NR}\neg\text{C}}$, respectively.

The fact that (41) is dead weight in Snakebite$^{\text{NRC}}$'s intuitive validity is highly relevant when we try to determine whether and how Snakebite$^{\text{NRC}}$ supports Transfer NRC. As Schnall and Widerker (2012) explain, the intuitive validity of an inference rule's instance supports the rule's validity by eliciting the logical and conceptual intuition that the rule is itself valid. However, since (41) is dead weight in Snakebite$^{\text{NRC}}$'s intuitive validity, Snakebite$^{\text{NRC}}$'s intuitive validity does *not* elicit any logical and conceptual intuitions concerning rules with a premise-scheme of which (41) is an instance. Specifically, Snakebite$^{\text{NRC}}$'s intuitive validity does not elicit the logical and conceptual intuition that Transfer NRC is valid. Instead, Snakebite$^{\text{NRC}}$'s intuitive validity can only elicit logical and conceptual intuitions supporting the validity of rules whose first premise-scheme has (39) as an instance, whose second premise-scheme has (40) as an instance, whose conclusion-scheme has (12) as an instance, and that is all. Some options as to what these rules might be are Capes' original rule B*, a generalized form of B* [say, $\text{NR}(C_p)$, $\text{NR}(C_p \supset p) \vdash \text{NR}(p)$], and the original Transfer NR rule.

Of course, even though Snakebite$^{\text{NRC}}$'s intuitive validity does not elicit the logical and conceptual intuition that Transfer NRC is valid, it might nonetheless support Transfer NRC's validity in some other way. For example, if Snakebite$^{\text{NRC}}$'s intuitive validity supports the validity of the original Transfer NR rule, then by monotonicity, it also supports Transfer NRC's validity *in an indirect way*. Besides, perhaps Snakebite$^{\text{NRC}}$'s intuitive validity supports Transfer NRC's validity in some other way entirely, which does not involve eliciting any of the intuitions Schnall and Widerker (2012) discuss.

In spite of this, the following general point clearly stands: Snakebite*'s intuitive validity plausibly elicits the intuition that B* is valid. Snakebite$^{\text{NRC}}$'s intuitive validity does not elicit the intuition that Transfer NRC is valid. So there is a distinctively direct way in which B* is—while Transfer NRC is not—supported by Snakebite-inspired instances. Consequently, B* is better supported by Snakebite-inspired instances than Transfer NRC. *A fortiori*, Snakebite* supports B* better than it supports Transfer NRC. It is thus false that Snakebite-inspired rule instances which support B* also support Transfer NRC at least as well. This removes the concern that B* does not pose a *prima facie* threat to compatibilism.

10 Conclusion

Two arguments suggest that Transfer NR does not pose even a *prima facie* threat to compatibilism. The first, due to McKenna, targets Transfer NR's original version and maintains that discussions of its validity "should never have even gotten off the ground" (2008, p. 370). Upon consideration, and *pace* Loewenstein, I find Schnall and Widerker's responses to this argument to be correct. But in discussing Loewenstein, I also find that there is a second way of arguing that Transfer NR does not pose a *prima facie* threat to compatibilism. This way of arguing is due to Fischer, and targets versions of Transfer NR that are not (at present, at least) subject to counterexamples. Given any such version, R, a Fischer-style argument requires showing that R-instances support some compatibilist-friendly R-surrogate at least as well as they support R. Fischer proposes Transfer NRC as this surrogate. However, further consideration suggests that when Capes' (2016) B* is substituted for R, we find that B*-instances do *not* support Transfer NRC at least as well as they support B*. So a Fischer-style argument that B* does not pose a *prima facie* threat to compatibilism would not be successful. Of course, the compatibilist might be able to provide some compatibilist-friendly B*-surrogate other than Transfer NRC. She might even be able to show that this surrogate is at least as well supported by B*-instances as B* is. But this is an open challenge. Until the compatibilist rises to it, we shall have every reason to hold that B* poses a *prima facie* threat to compatibilism.

References

Capes, J. A. (2016). Incompatibilism and the transfer of non-responsibility. *Philosophical Studies*, *173*, 1477–1495.

Cyr, T. W. (forthcoming). Semicompatibilism and moral responsibility for actions and omissions: In defence of symmetrical requirements. *Australasian Journal of Philosophy*. doi:10.1080/00048402.2020.1738512

Fischer, J. M. (2004). The transfer of nonresponsibility. In J. K. Campbell, M. O'Rourke, & D. Shier (Eds.), *Freedom and determinism* (Chap. 9, pp. 189–209). Cambridge, MA: MIT Press.

Fischer, J. M. (2006). The transfer of nonresponsibility. In *My way - essays on moral responsibility* (Chap. 8, pp. 159–174). Oxford University Press.

Fischer, J. M., & Ravizza, M. (1998). *Responsibility and control: A theory of moral responsibility*. Cambridge: Cambridge University Press.

Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *Journal of Philosophy*, *66*, 829–839.

Haji, I. (2008). Reflections on the incompatibilist's direct argument. *Erkenntnis*, *68*, 1–19.

Haji, I. (2010). On the direct argument for the incompatibility of determinism and moral responsibility. *Grazer Philosophische Studien*, *80*, 111–130.

Hermes, C. (2014a). A counterexample to A. *Philosophia*, *42*, 387–389.

Hermes, C. (2014b). Truthmakers and the direct argument. *Philosophical Studies*, *167*, 401–418.

Kearns, S. (2011). Responsibility for necessities. *Philosophical Studies*, *155*, 307–324.

Loewenstein, Y. (2016). Why the direct argument does not shift the burden of proof. *Journal of Philosophy*, *113*, 210–223.

McKenna, M. (2001). Source incompatibilism, ultimacy, and the transfer of non-responsibility. *American Philosophical Quarterly*, *38*, 37–51.

McKenna, M. (2008). Saying good-bye to the direct argument the right way. *Philosophical Review*, *117*, 349–383.

Ravizza, M. (1994). Semi-compatibilism and the transfer of non-responsibility. *Philosophical Studies*, *75*, 61–93.

Robinson, M. (2016). Truthmakers, moral responsibility, and an alleged counterexample to rule a. *Erkenntnis*, *81*, 1333–1339.

Schnall, I. M., & Widerker, D. (2012). The direct argument and the burden of proof. *Analysis*, *72*, 25–36.

Shabo, S. (2010a). Against logical versions of the direct argument: A new counterexample. *American Philosophical Quarterly*, *47*, 239–252.

Shabo, S. (2010b). The fate of the direct argument and the case for incompatibilism. *Philosophical Studies*, *150*, 405–424.

Singer, P. (1972). Famine, affluence, and morality. *Philosophy & Public Affairs*, *1*, 229–243.

Stump, E. (2000). The direct argument for incompatibilism. *Philosophy and Phenomenological Research*, *61*, 459–466.

Stump, E. (2002). Control and causal determinism. In S. Buss & L. Overton (Eds.), *Contours of agency: Essays on themes from harry frankfurt* (Chap. 2, pp. 33–60). Cambridge, MA: MIT Press.

Stump, E., & Fischer, J. M. (2000). Transfer principles and moral responsibility. *Philosophical Perspectives*, *14*, 47–55.

Tucker, C. (Ed.). (2013). *Seemings and justification*. Oxford: Oxford University Press.

Turner, P. R., & Capes, J. (2018). Rule A. *Pacific Philosophical Quarterly*, *99*, 580–595.

van Inwagen, P. (1983). *An essay on free will*. Oxford: Oxford University Press.

Warfield, T. A. (1996). Determinism and moral responsibility are incompatible. *Philosophical Topics*, *24*, 215–226.

Widerker, D. (2002). Farewell to the direct argument. *Journal of Philosophy*, *99*, 316–324.

Widerker, D., & Schnall, I. M. (2014). The direct argument for incompatibilism. In D. Palmer (Ed.), *Libertarian free will: Contemporary debates* (Chap. 7, pp. 88–106). Oxford: Oxford University Press.