The Zygote Argument is invalid: Now what?

Kristin Mickelson
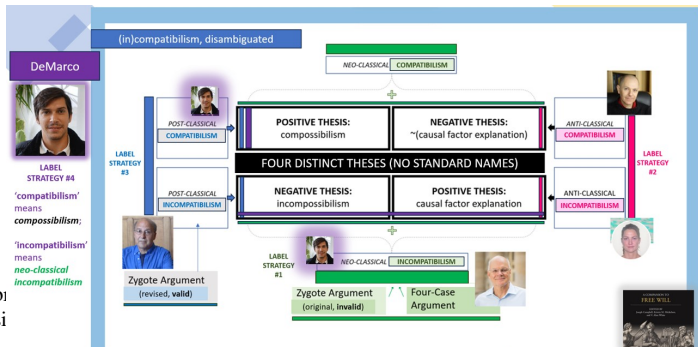
Abstract Alfred Mele's original Zygote Argument is invalid. At most, its premises entail the negative thesis that free action is *incompossible* with deterministic laws, but its conclusion asserts the positive thesis that deterministic laws *preclude* (make impossible, undermine) free action. The original, explanatory conclusion of the Zygote Argument can be defended only by supplementing it with a best-explanation argument that identifies deterministic laws as menacing. (By the same reasoning, it follows that every manipulation argument pinpointing a specific threat to free will requires a best-explanation argument). Arguably, though, the best explanation for the manipulation victim's lack of freedom and responsibility is his *constitutive luck*, which is a problem irrespective of the natural laws that obtain. This proposed explanation leads to a new "diagnostic" version of the Zygote Argument which concludes that free action is impossible even though deterministic laws pose no threat whatsoever to free will.

Keywords Free will · Zygote Argument · Manipulation · Best-explanation argument · Constitutive luck · Incompossibilism

1 Introduction

Alfred Mele's original Zygote Argument (e.g. 2006, 2008) is one of the most familiar arguments in the contemporary free-will debate. The argument is also invalid. The argument's conclusion is a statement of *incompatibilism*, roughly the positive view that deterministic laws undermine free agency, but its premises support mere *incompossibilism*, the negative view that free will and deterministic laws are *incompossible* phenomena (i.e. they cannot possibly co-exist or co-obtain).

K. Mickelson (✉)
University of Minnesota at Morris, 600 E. 4th Street, Morris, MN 56267, USA
e-mail: mickelkm@morris.umn.edu

As such, there are two ways that one might repair the argument. The simplest repair strategy is to weaken the original conclusion so that the Zygote Argument concludes to mere incompossibilism. However, if the original, explanatory conclusion of Mele's argument is to be defended, the Zygote Argument must be amended to include a premise that identifies deterministic laws as a freedom-undermining feature of the manipulation story.

Since opponents of the Zygote Argument cannot be expected to accept a premise that pinpoints deterministic laws as a specific threat to free will, the new premise requires a defense. Using another, logically independent argument for incompatibilism to defend this premise would reduce the manipulation argument to a façade for that other argument. As such, it seems the only way to provide a positive defense of the new premise is to give a best-explanation argument. By the same reasoning, *any* manipulation argument that has a premise (and conclusion) that pinpoints a specific threat to free will must include a best-explanation argument. This is noteworthy because Mele's zygote story may be used to identify something other than deterministic laws as a threat to free will. For instance, one might reasonably argue that the best explanation for the manipulation victim's lack of free will is that he did not self-create in the way required to satisfy a sourcehood requirement for free and responsible action—and his failure to self-create in the requisite way is *not* due to his being subject to deterministic natural laws. Developing the Zygote Argument along these lines would result in a version of the Zygote Argument that concludes to incompossibilism, but not incompatibilism. Indeed, assuming standard possible worlds semantics, this new version of the Zygote Argument would be an argument *against* the historically significant view that deterministic laws, if they were to obtain, would undermine free will.

This essay begins, in Sect. 2, with a review and critique of the original Zygote Argument and two alternative versions of the argument that Mele has recently developed. None of these arguments constitutes a valid defense of the incompatibilist view that someone who is subject to deterministic laws cannot act freely (at least in part) *because* she is subject to deterministic laws. In order to illuminate the logical structure of a manipulation argument that does conclude to incompatibilism, I introduce a new formal template that represents the logical structure of "diagnostic" manipulation arguments. Diagnostic manipulation arguments differ from standard manipulation arguments primarily in virtue of having a "diagnostic premise," i.e. a premise that gives a diagnosis of the freedom-undermining feature(s) in the manipulation story. In Sect. 3, I urge that there is only one viable way to defend the diagnostic premise of a diagnostic manipulation argument: a best-explanation argument. The conclusion that any manipulation argument for incompatibilism requires a best-explanation argument rests upon the assumption that a philosopher may coherently hold that free will is *incompossible* with deterministic laws while denying that deterministic laws pose a *threat* to free will. Based on a brief discussion of Galen Strawson's Basic Argument, I confirm in Sect. 4 that an argument for incompossibilism need not constitute an argument for incompatibilism. I then outline a new diagnostic version of the Zygote Argument in Sect. 5. According to this new Zygote Argument, the manipulation victim in Mele's zygote story lacks free will because he suffers from freedom-undermining

constitutive luck. That is, the manipulation victim has failed to self-create in a freedom-relevant way, where his failure to self-create in the requisite way is *not* due to his being subject to deterministic natural laws. If this explanation is right, then the lesson of Mele's zygote story is—at least when we assume standard possible worlds semantics for modal claims—that free-will impossibilism (a.k.a. skepticism) and incompossibilism are true, but *incompatibilism is false*.

2 The original Zygote Argument is invalid

Mele's original Zygote Argument is based on a story in which the goddess Diana creates a zygote that ultimately grows into a human adult, Ernie. The details of Mele's story are consistent with various modal interpretations, but the gist of the story is this: Diana wants a particular event E to occur at time t^* in deterministic universe U, so she performs a single creative act at t that ensures that E will eventually come to pass at t^*. Specifically, Diana creates zygote Z at time t and thereby makes it inevitable (given the state of U at t and that the laws of nature at U are deterministic) that Z develops into Ernie, and that Ernie performs an action A that brings about E at t^*.

Generalizing from this zygote story to a normal deterministic scenario, Mele develops an argument which he formally summarizes as follows, henceforth "ZA":

1. Because of the way his zygote was produced in his deterministic universe, Ernie is not a free agent and is not morally responsible for anything.
2. Concerning free action and moral responsibility of the beings into whom the zygotes develop, there is no significant difference between the way Ernie's zygote comes to exist and the way any normal human zygote comes to exist in a deterministic universe.
3. So determinism *precludes* free action and moral responsibility (Mele 2006: 189, 2008: 280; my emphasis).

Given the standard English definition of 'preclude,' the conclusion of ZA asserts that there is something about deterministic laws in virtue of which they *undermine* free and morally-responsible agency.

However, a more careful review of the premises of ZA reveal that neither premise 1 nor 2 identifies deterministic laws as "menacing," i.e. freedom- and responsibility-undermining.[1] Yes, Ernie lives in a deterministic universe, but premise 1 does not assert that it is *in virtue of* being subject to deterministic laws that Ernie lacks free will. For all that is said in the premises of ZA, the fact that Ernie's universe is deterministic may or may not be relevant to his status as a free agent. If anything, the emphasis in premise 1 of ZA is on the way that Ernie's zygote was *produced*. However, Mele's description of ZA makes it clear that we should not read premise 1 as promoting a particular account of Ernie's lack of freedom and responsibility. Mele explicitly *denies* that his argument includes a

---

[1] I borrow the convenient term "menacing" from Ishtiyaque Haji.

"best-explanation premise" that forwards an explanation of Ernie's lack of freedom and responsibility (2008: 286), and specifically says that ZA has *no* premise that "zeroes in on determinism" as a specific threat to freedom and responsibility (2008: 284). In the most recent formal summaries of the Zygote Argument, Mele has dropped the "because" clause from premise 1 altogether (see discussions of ZAM-1 and ZAM-2 below). As such, it seems that the "because" clause in premise 1 is meant to point us in the general direction of (what seems to be) the source of Ernie's problems, but does not positively identify deterministic causation as menacing. So, at best, what follows from the non-explanatory premises of ZA is that free action and moral responsibility are incompossible with deterministic laws. That is, ZA's premises do not entail the explanatory thesis that deterministic laws *preclude*—make impossible, undermine—free action and moral responsibility. In short, ZA is invalid.

Of course, ZA is only a formal summary of an argument. As such, one might wonder whether the problems with ZA arise because it does not adequately represent the natural-language version of the Zygote Argument. However, a closer look at the latter reveals that Mele does not pinpoint deterministic laws as a threat to freedom and/or moral responsibility in the natural-language versions of his argument either. Regarding premise 1, Mele explains that "Premise 1 of the zygote argument is an assertion about a case" and "the zygote argument is supposed to use an intuition about Ernie as a step toward the conclusion that incompatibilism is true" (2006: 192). The test of the truth of premise 1 is an intuitive judgment about Ernie, namely that he lacks freedom and responsibility for his actions.[2] Whether or not a rational person would have this intuition, what we might call a "victim intuition," is the subject of much debate—notably, Mele says he does not have it (e.g., 2013: 183). What is crucial here, though, is that Mele does not argue that a victim intuition is a specific response, let alone a *rational* response, to the *deterministic causation* in Ernie's story. In fact, Mele tries to cast doubt on the proposal that a victim intuition tracks deterministic causation by considering some intuitive reactions to *indeterministic* versions of Ernie's story (e.g., 2013: 177). It is also worth noting that Mele uses the same general strategy to critique Pereboom's "Four-case Argument". According to Mele (e.g. 2006: 144), the Four-case Argument is unsuccessful primarily because Pereboom provides inadequate support for his (empirical) claim that *deterministic causation* drives the intuition that the manipulation lacks freedom and responsibility and for his (metaphysical) claim that

---

[2] In early discussions of the Zygote Argument, Mele allows that a proponent of ZA could also give a positive argument in defense of premise 1. Citing a passage from Thomas Kapitan (2000: 90), Mele supplies one "predictable" defense of premise 1, namely that Ernie is "deliberately caused to behave in a certain way in much the same way that designers of robots program the responses of their machines to various stimuli" (Mele 2006: 189, 2008: 280). Notably, this argument (if it deserves to be called one) mentions but *does not identify* deterministic laws or deterministic causation as a threat to free will. (It is unclear, for example, whether the menacing feature of Ernie's story is that he was *programmed*, that he was *deliberately programmed,* that he was *deliberately programmed and then subjected to deterministic laws*, or some other combination of the features that Mele mentions). However, in more recent (e.g. Mele 2013) formulations of the zygote argument, Mele seems to have adopted the more standard line according to which the basis for one's judgment of the truth of premise 1 rests *entirely* on one's intuitive judgment of Ernie's status as a free and responsible agent.

the manipulation victims lack freedom and responsibility (in part) because they were subject to deterministic causation. By contrast, Mele considers the Zygote Argument to be a highly compelling argument that is "a significant part of what prevents some of us [including Mele] from coming down off the fence and endorsing compatibilism" (2006: 192). Clearly, Mele does not think that the Zygote Argument is subject to the criticisms that he launches against Pereboom's Four-case Argument. This lends further support to the conclusion that Mele does not intend for the "because" clause in premise 1 to be interpreted as a best-explanation claim, and confirms that Mele does not take up the burden of defending the best-explanation claim that deterministic laws are menacing in the course of presenting the Zygote Argument.

Although deterministic laws are present in both the manipulation and determination scenarios, Mele also makes it clear that he does not identify *any* particular threat to freedom or responsibility in the course of defending premise 2 of ZA. Mele points out that, for all he has said in defense of premise 2, there may be *no* freedom- or responsibility-undermining feature in either case:

> Although the argument I have sketched for premise 2 sounds a bit like the consequence argument, it is significantly different. The consequence argument is an argument for incompatibilism. The argument for premise 2 is, by itself, consistent with compatibilism. The thesis that the cross-universe difference in what caused Z [the zygote that becomes Ernie] does not support any cross-universe difference in freedom or moral responsibility is consistent with Ernie's acting freely and morally responsibly in both universes, as is premise 2. (Mele 2006: 190)[3]

Mele provides a strictly negative "no-difference" defense of premise 2 in which he argues that none of the differences between his zygote story and the normal determination scenario is a freedom- or responsibility-relevant difference. Mele does not, in addition, defend premise 2 by identifying some freedom-undermining feature that is common to both scenarios. Using this no-difference defense strategy, Mele does just enough to shift the burden of proof to the opponent of his argument: perhaps Mele has overlooked some difference between the manipulation and normal determination scenarios that justifies rejecting premise 2, but it falls on the opponent of the argument to identify it.

In sum, Mele offers no defense of the claim that deterministic causation poses a threat to free will and moral responsibility in the course of presenting the original Zygote Argument. Indeed, Mele does not even suggest that deterministic causation or causal laws of any sort are worth mentioning in relation to freedom and responsibility. Thus, the premises of ZA do not undersell the premises of the

---

[3] Notably, the meaning of Mele's term 'incompatibilism' is unclear in this context. He states that incompatibilism is "The thesis that neither free action nor moral responsibility is *compatible* with the truth of determinism" (2006; my emphasis), but does not adequately specify the notion of "(in)compatibility" at issue.

natural-language Zygote Argument. The most that follows from Mele's argument is the negative thesis that no one can perform a free action when subject to deterministic laws. So, contrary to the incompatibilist conclusion that Mele draws from his argument, i.e. the conclusion of ZA, the original Zygote Argument does not constitute a defense of the incompatibilist claim that deterministic laws *preclude* freedom or moral responsibility. Even in its full, natural-language form, the original Zygote Argument is invalid.

Others (e.g., Kearns 2012) have pointed out that ZA purports to be an argument for the incompatibilist thesis that deterministic laws *on their own* preclude free will when the argument may best be understood as a defense of a more modest view, such as: necessarily, if it is causally determined that someone S performs an action A, then S does not freely perform A *only in part* because S is causally determined to A. Partly in response to such concerns, Mele has offered to new version of the Zygote Argument that he calls "ZAM." In fact, Mele has offered two versions of ZAM which differ slightly in their respective conclusions, but neither has the same conclusion as ZA.[4] The first version of ZAM, henceforth "ZAM-1", goes as follows:

1. Ernie is not morally responsible for anything he does.
2. Concerning moral responsibility of the beings into whom the zygotes develop, there is no significant difference between the way Ernie's zygote comes to exist and the way any normal human zygote comes to exist in a deterministic universe.
3. So determinism *precludes* moral responsibility—at least for human beings who develop from normal human zygotes. (Mele 2012; my emphasis)

Notably, the conclusion of ZAM-1 is more modest than the conclusion of ZA in two ways. First, ZAM-1 does not address the issue of *free* action, but only the moral responsibility for one's actions. Second, and more importantly for our purposes, ZAM-1's conclusion is explicitly restricted to certain sorts of beings with a certain sort of history, namely *human beings* who *develop from normal human zygotes*. In adding these restrictions, Mele acknowledges that *being human* or *being born as a zygote* may be freedom- or responsibility-relevant features without asserting that they are. The rhetorical advantage of adding these restrictions is clear: even when they are added, the conclusion of ZAM-1 covers the actions of beings that we most care about—*us*—yet Mele avoids getting bogged down in a tricky debate about which specific features are freedom- and responsibility-relevant ones.[5] While these differences are worth noting,

---

[4] The creation stories grounding ZA and ZAM are slightly different. These differences may be relevant to the soundness of the argument, but the differences are not relevant to the present discussion of the logical properties shared by these arguments. As such, I will not address their differences here.

[5] On the downside, adding such restrictions even when it is unclear that they are relevant to the existence of freedom or responsibility may obscure the logic and lesson of an argument (e.g., see Campbell's (2007) "No-past Objection" to the Consequence Argument). Mele's repetition of the fact that Ernie lives in a deterministic universe gives rise to the misimpression that the argument pinpoints deterministic laws as a threat to free will (at least for beings of a certain sort), which may partially explain why the invalidity of ZA was overlooked for so long. The arguments in Sect. 5 of this essay suggest that the restrictions that Mele's builds into the conclusion of ZAM—including "in a deterministic universe"—are *not* freedom- or responsibility-relevant ones.

the conclusions of ZA and ZAM-2 are also the same in one crucial way: each positively identifies deterministic laws as a *threat* to moral responsibility. However, like ZA, the premises of ZAM-1 do not state or entail any particular diagnosis of what deprives (someone like) Ernie of moral responsibility. So, like ZA, ZAM-1 is invalid.

One might be tempted to brush off my complaints given that the Zygote Argument could easily be reformulated as a valid argument. In particular, one might suggest that we formulate the Zygote Argument as an instance of an argument template that Mele calls the "Straight Manipulation Argument" (SMA):

1. (Manny Premise) Manny does not freely A and is not morally responsible for A-ing.
2. (No-difference Premise) Concerning free action and moral responsibility, there is no significant difference between Manny's A-ing and any candidate for a free and morally responsible action in a deterministic universe.
3. So no candidate for a free and morally responsible action in a deterministic universe is a free action nor an action for which its agent is morally responsible[...]. (2008: 265)[6]

Looking at the premises of ZA, we can see that the Zygote Argument provides the requisite content for both a "Manny Premise" and a "No-difference Premise." Thus, there seems to be no formal problem with formulating the Zygote Argument as an instance of SMA.[7]

---

[6] In my view, the formal structure of this template (as well as McKenna's preferred statement of the general template for manipulation arguments (e.g., McKenna 2012: 151)) is slightly misleading. Non-diagnostic manipulation arguments of this sort have a simple modus ponens form: *The Non-diagnostic Manipulation Argument Template:*

1. *Victim Premise:* Due to some feature of the (apparent) manipulation scenario, the (apparent) manipulation victim *S* is not free or responsible for performing an action *A*.
2. *Generalization Premise*: If *S* is not free or responsible for performing *A*, then then no one in a (normal) determination scenario is free or responsible.
3. *Conclusion:* No one is free or responsible in the determination scenario.

Premise 2 of Mele's SMA is really a supporting claim for the Generalization Premise of the above template. In support of the truth of the Generalization Premise, a philosopher might point to such things as a positive diagnosis the menacing feature that is common to both scenarios and the methodological principle that like cases must be judged alike.

[7] Mele (2008) does not consider the Zygote Argument to be an instance of SMA. However, Mele does not explicitly deny that the Zygote Argument has the *logical form* outlined by SMA. Rather, Mele argues that because the "original design" story he tells does not involve any *genuine* manipulation, the Zygote Argument is not technically a *manipulation* argument at all. In my view, Mele's proposed individuation principle gives rise to an overly narrow conception of the manipulation-argument strategy. There certainly *seems* to be some sort of manipulation taking place when Diana creates Ernie as she does, and this *appearance* of manipulation seems to be essential to the common intuitive reaction to the case (but the *genuineness* of the manipulation does not). Moreover, some (e.g., Barnes 2013) have argued that there *is* genuine manipulation in Mele's zygote case. If there is genuine manipulation in the case, then Mele has categorized the Zygote Argument wrongly according to his own individuation principle. For reasons such as these, I contend that the Zygote Argument is best classified as a manipulation argument. That said, manipulation arguments that rest on "original-design" stories are surely worth distinguishing from others on the grounds that they are the most compelling instances of the manipulation-argument strategy.

Looking at the conclusion of SMA, though, we see the downside of this retooling strategy. The conclusion of ZA entails the conclusion of SMA, but not vice versa. The conclusion of SMA is that "no candidate for a free and morally responsible action in a deterministic universe is a free action nor an action for which its agent is morally responsible." The conclusion of SMA does not indicate *why* there are no free agents in any deterministic universe; it does not tell us *in virtue of what* there are no free agents living in any deterministic universe. More specifically, the conclusion does not assert that *being subject to deterministic laws* [either alone or in conjunction with other feature(s) of the world] ever deprives people of their free will. Since SMA does not outline an argument for the positive view that deterministic laws pose a threat to free will, we cannot reformulate the Zygote Argument as an instance of SMA without sacrificing ZA's original, explanatory conclusion.[8]

This brings us to Mele's most recent version of the Zygote Argument, the second version of ZAM, henceforth "ZAM-2". ZAM-2 is, roughly, an instance of SMA. ZAM-2 has the same premises as ZAM-1 (above), but concludes to: "So in no possible deterministic world in which a human being develops from a normal human zygote is that human being morally responsible for anything he or she does (Mele 2013: 176)". Like ZAM-1, the conclusion of ZAM-2 is restricted to the moral responsibility of humans with a certain sort of past. Unlike ZAM-1, ZAM-2 does not specify which, if any, of the restrictions on its conclusion—having a normal-zygote history, being human, etc.—is relevant to free will and/or moral responsibility. As such, Mele's shift from ZAM-1 to ZAM-2 has come at a cost: ZAM-2 is valid, but its conclusion does not pinpoint deterministic laws as a specific threat to (free will or, thereby, to) moral responsibility.

Having seen that the conclusions of ZA and ZAM-1 are substantively different from the conclusions of SMA and ZAM-2, the reader may wonder why each is universally described as an argument for incompatibilism. The term 'incompatibilism' was originally used to name (roughly) the view that necessarily, if determinism is true, then the free-will thesis is false.[9] Notably, this conditional claim may be true even if the truth of determinism is not relevant to the falsity of the free-will thesis; it expresses only the negative thesis that the conjunction of determinism and the free-will thesis is necessarily false. Restated in first-order language, incompatibilism was originally characterized as the non-explanatory view that deterministic laws and free will cannot co-exist, i.e. these phenomena are incompossible. The respective

---

[8] This observation about SMA is noteworthy because SMA is (as far as I can tell) formally equivalent to Michael McKenna's generic template for manipulation arguments, "The Manipulation Argument" (e.g. 2012:151). As such, McKenna's template, too, fails to outline an argument for incompatibilism. Thus, recasting the Zygote Argument as a (valid) instance of McKenna's template would not be relevantly different from recasting it as an instance of SMA.

[9] It seems that the terms 'compatibilist' and 'incompatibilist' were introduced by Keith Lehrer, and were first used in print in Leher's (1960) dissertation; he introduced the more standard characterizations of these terms in print eight years later (Cornman and Lehrer 1968: 130). It seems that the corresponding terms 'compatibilism' and 'incompatibilism' were first used in print by van Inwagen in his (1969) dissertation, but van Inwagen (in correspondence) credits Lehrer–who was Second Reader on van Inwagen's dissertation defense committee–with the coining of these terms as well.

conclusions of SMA, ZA, ZAM-1, and ZAM-2 each qualifies as a statement of this *non-explanatory* incompatibilism. However, some contemporary philosophers use 'incompatibilism' more narrowly to name the positive view that someone who is subject to deterministic laws would lack free will *because* she is subject to deterministic laws (e.g., McKenna 2010: 432; Vihvelin 2013: 24). The conclusions of ZA and ZAM-1 are statements of this *explanatory* incompatibilism, but the conclusions of SMA and ZAM-2 are not.

In order to make it easier to track the distinction between these explanatory and non-explanatory views, I will hereafter use 'incompatibilism' to refer only to the positive, explanatory view that necessarily, anyone who is subject to deterministic laws is unfree (at least in part) because or in virtue of being subject to such laws; 'compatibilism' will name the negation of incompatibilism. I will use the apt term 'incompossibilism' to name the negative, non-explanatory view that deterministic laws and free action are incompossible, and 'compossibilism' to name the contradictory of incompossibilism. As I have characterized these views, incompatibilism entails incompossibilism, but incompossibilism does not entail incompatibilism.[10] Using this terminology, we can easily describe the key difference between the conclusions of ZA and ZAM-2: ZAM-2's conclusion is a statement of mere incompossibilism while ZA's conclusion is a statement of incompatibilism. We can also easily summarize the formal problem with ZA and ZAM-1: each argument is invalid because its premises entail mere incompossibilism, yet each concludes to incompatibilism.

Whatever might be said about the relative merits of using a manipulation argument to defend mere incompossibilism as opposed to incompatibilism, this much should now be clear: a manipulation argument does not constitute a defense of incompatibilism unless the premises (and not just the conclusion) of the manipulation argument "zero in" on determination.

3 Manipulation arguments and the importance of best-explanation arguments

This brings us to a question that has been raised in various ways in current literature: Does a successful manipulation argument require a best-explanation argument? With the incompossibilism/incompatibilism distinction in hand, we can see that there is no single correct answer to this question. Rather, the answer depends upon what one takes the goal of a manipulation argument to be.

If one thinks that a "successful" manipulation argument would be a persuasive argument that concludes to mere incompossibilism, then a successful manipulation

---

[10] Technically, incompatibilism entails only the *qualified* incompossibilist thesis that there is no possible universe in which deterministic laws obtain and *someone who is subject to the laws* performs a free action. According to *unqualified* incompossibilism, it is even impossible for a being who is not subject to the natural laws (e.g., some god-like being who could change or violate the laws of nature) to act freely in a universe with deterministic laws. Since the differences between incompossibilism and more qualified versions of incompossibilism are not pressing in the present context, I will ignore these subtleties for the purposes of this essay.

argument does not require a best-explanation argument that pinpoints the menacing feature of the manipulation scenario. As discussed above, it seems that an adequate defense of an instance of the No-difference Premise of SMA can be given in completely negative terms. Simply by ruling out all apparent freedom- and responsibility-relevant differences between the manipulation and determination cases, the proponent of a manipulation argument shifts the burden of proof to those who are targeted by the argument: in order to deny the truth of the No-difference Premise, a freedom- or responsibility-relevant difference must be identified. So, it seems that any proponent of compossibilism who has a strong *victim intuition* must either (1) defend a "soft-line" reply to the manipulation argument by identifying a principled difference between the manipulation and natural determination scenarios, or (2) defend a "hard-line" reply to the argument by rejecting premise 1, (e.g., by presenting a compelling error-theory for her victim intuition). As such, even a relatively modest argument for incompossibilism, such as ZAM-2, presents a significant challenge to defenders of compossibilism.[11]

On the other hand, if a philosopher wishes—as the incompatibilist does—to use a manipulation argument to support the conclusion that some particular feature of a world (e.g., causal determination, method of creation, etc.) is menacing, then she must appeal to a best-explanation argument. In order to understand why, let us consider the formal features that would be required of a manipulation argument for incompatibilism—something like:

The Diagnostic Manipulation Argument Template

1. *Victim Premise:* Due to some feature of the (apparent) manipulation scenario, the (apparent) manipulation victim S is not free or responsible for performing an action A.
2. *Diagnostic Premise*: The menacing feature of the (apparent) manipulation scenario—the feature that accounts for S's lack of freedom and responsibility—is F.
3. *Same-Feature Premise*: F is present in both the (apparent) manipulation scenario and any normal determination scenario.
4. *Conclusion:* No one is free or responsible in any normal determination scenario in virtue of F.

The proponent of a diagnostic manipulation argument must develop a best-explanation argument to defend the Diagnostic Premise of an instance of the

---

[11] I am oversimplifying the dialectic here. I do not mean to imply that a manipulation argument presents a challenge only to those philosophers who have a (strong) victim intuition in response to the argument's manipulation story. It seems that a manipulation argument will have at least some purchase so long as there is *a difference* between one's intuitive judgments of the manipulation victim and the naturally determined agent, whether in *degree* (e.g. one has a strong intuition that the naturally determined agent is free but only a weak intuition that the manipulation victim is free) or in *kind* (e.g., one has a victim intuition in response to the manipulation case but also the intuition that the merely determined agent is free). However interesting, these complicated dialectical issues are not pressing in the present context, where the primary focus is exposing the logical structures of manipulation arguments and not identifying the full range of their target audience.

Diagnostic Manipulation Argument template, I contend, because there is no other viable way to defend it.

Let us consider the resources at the disposal of the incompatibilist trying to defend the (highly contentious) diagnostic claim that Ernie's being subject to deterministic laws—meaning that the laws account for the diachronic evolution of Ernie's states and he is unable to *violate* or *change* the natural laws—is what makes Ernie unfree. The incompatibilist might try to defend this claim by appealing to another, logically independent argument for incompatibilism, such as the Consequence Argument.[12] The general strategy of using a free-standing argument in defense of a Diagnostic Premise would be unfruitful for at least two reasons. First, manipulation arguments are designed to be on a par with more traditional arguments for incompatibilism—indeed, they are so attractive precisely because they promise to succeed where these other arguments have failed. This status cannot be maintained if the linchpin of the manipulation is argument just *is* some other (new or old) argument for incompatibilism. Second, if only a stand-alone argument for incompatibilism were used to defend the Diagnostic Premise, then the manipulation argument would not work as designed—namely, as a *persuasive* argument— because the victim intuition (i.e. the intuition that the manipulation victim is not free or responsible) would not play an essential role in the argument. That is, one's having a victim intuition would no longer set the hook by which one is then dragged to the conclusion that incompatibilism is true. Thus, if the manipulation argument is to be more than a façade for some other argument, the proponent of a diagnostic manipulation argument for incompatibilism must look elsewhere for a defense of its diagnostic premise.

A best-explanation fits the bill. First, a best-explanation argument defending the Diagnostic Premise must appeal to the details of the manipulation story, for the goal of the best-explanation argument is to identify the menacing feature F in that story. For example, Ernie's lack of freedom cannot be explained without appealing to specific features of Ernie's story—the deterministic laws, Diana's foreknowledge, etc.—and identifying which of these account for Ernie's lack of free will and moral responsibility. Second, a best-explanation defense of the Diagnostic Premise constitutes a positive defense of the rationality of the victim intuition (i.e. it is a rational response to F), thereby increasing the plausibility of the Victim Premise. That is, in offering a best-explanation of Ernie's lack of moral responsibility, one also provides a defense of the rationality of the intuition that Ernie is neither free nor responsible.[13] It seems, then, that a best-explanation defense of the Diagnostic

---

[12] Notably, as van Inwagen defines 'incompatibilism', incompatibilism is the mere denial of the thesis that determinism and the free-will thesis could both be true (e.g. 2008: 330). Standard formulations of the Consequence Argument defend only this negative, inconsistency thesis and *not* the positive, explanatory thesis that (necessarily) if determinism is true, then the free-will thesis is false *because determinism is true*. As such, there is room to argue–as I do in "No Past? No Problem" (unpublished manuscript)–that the Consequence Argument itself would need to be supplemented with a best-explanation argument in order to pinpoint a specific threat to free will, and there is even room to deny that the lesson of the Consequence Argument (if sound) is that being subject to deterministic laws undermines free will.

[13] One might argue that diagnostic manipulation arguments are not as purely "persuasive" as non-diagnostic manipulation arguments. Because the best-explanation argument defense of the Diagnostic

Premise does just what is needed from such a defense: it defends the diagnostic claim in a way that makes the manipulation story, the victim intuition, and the Victim Premise essential parts of the overall argument.

What other than a best-explanation argument would establish the crucial connection between the victim intuition and the determination in the manipulation story without constituting a logically- independent argument for incompatibilism? I see no alternative. If there is no alternative, then a best-explanation argument will be an essential part of the defense of every diagnostic manipulation argument, not just incompatibilist-promoting versions of the Zygote Argument.[14]

4 Incompossibilism without incompatibilism

As a technical matter, generating a best-explanation argument that identifies some feature of a manipulation story as menacing is not a difficult task. Obviously, though, adding a best-explanation argument to a manipulation argument would make the latter vulnerable to new attacks. That being so, the reader might wonder if it is possible to avoid my conclusion that a best-explanation argument is a necessary part of any diagnostic manipulation argument by rejecting the incompatibilism/ incompossibilism distinction upon which the above critique of the Zygote Argument is based. Roughly, one might argue that in order for there to be a "metaphysically significant" difference between incompossibilism and incompatibilism, there would have to be some metaphysical view that implies the truth of incompossibilism but not the truth of incompatibilism—but there is no such view. If this line of thinking were right, then a best-explanation argument would not be required to upgrade the

---

Footnote 13 continued

Premise picks out a feature of the manipulation story which would make a victim intuition a rational response to the story, it also implies that anyone who fails to have the victim intuition is suffering from some sort of problem, e.g., irrationality or what Mele (2013: 182) has called "intuition deficit disorder." As such, rejecting the truth of the Victim Premise of a diagnostic manipulation argument (i.e. defending a "hard-line" reply) may be dialectically more complicated than rejecting the same claim in the context of a non-diagnostic manipulation argument. Such differences between the dialectics surrounding diagnostic and non-diagnostic manipulation arguments are worth exploring, but I gloss over these differences here because they are not relevant to my argument that a valid manipulation argument for incompatibilism must include a best-explanation argument.

[14] The reader should not take this conclusion as implying that all manipulation arguments for incompatibilism are instances of Mele's "Best-explanation Manipulation Argument" template (Mele 2008: 276). Mele's Best-explanation Manipulation Argument template mischaracterizes the role played by best-explanation arguments in manipulation arguments for incompatibilism. Mele describes a best-explanation manipulation argument as one that differs from an instance of SMA only in its second premise: the former has a "Best-explanation Premise" identifying the menacing feature of the manipulation scenario instead of the latter's No-difference Premise. As such, Mele's Best-explanation template does not make the case that the *same* menacing feature is present in both the manipulation scenario and the normal deterministic scenario. This is important because even when the menacing feature is correctly identified as a type of determination, it remains an open question whether the same menacing type of determination is present in the normal scenario because not all deterministic manipulation perfectly mimics natural determination cf. (Demetriou 2010). In order to bridge that gap, a manipulation argument needs something like what I (above) call a "Same-feature Premise," and Mele provides no such bridge premise in his template.

original Zygote Argument to an argument for incompatibilism. Rather, the incompatibilist would need only present the Zygote Argument against the *background* of a proper understanding of the shared metaphysical commitments of the incompossibilist and incompatibilist to show that Mele does get the incompatibilist conclusion "for free" from ZAM-2, an argument for mere incompossibilism.

However, incompossibilism and incompatibilism do come apart, and in interesting and unexplored ways. To illustrate, consider Galen Strawson's Basic Argument (1986, 1994, 2002, 2008). The Basic Argument is an argument for free-will *impossibilism* (a.k.a. skepticism), the view that it is metaphysically impossible for someone to perform a free action. The upshot of the argument is that there is a necessary sourcehood condition for free action that is impossible for anyone (even God) to satisfy. As G. Strawson explains, in order for a person to act freely, "there has to be, and cannot be, a *starting point* in the series of acts of [intentionally] bringing it about that one has a certain nature–a *starting point* that constitutes an act of ultimate self-origination" (Strawson 1998/2011; my emphasis). Anyone who does not satisfy this starting-point sourcehood condition suffers from a type of freedom- and responsibility-undermining *luck* (e.g. Strawson 1998/2011) commonly known as *constitutive luck*. It seems that such a stringent sourcehood condition could only be satisfied by a *causa sui*, i.e. someone who self-creates ex nihilo. Because such radical self-creation is impossible (the very notion of a *causa sui* is self-contradictory), so is free and responsible action.

Notably, the Basic Argument does not conclude that the natural laws which govern the evolution of someone S's universe—whether deterministic or indeterministic—*make it the case* or otherwise *account for the fact* that S is not a *causa sui* at any time *t* at which S exists. Rather, the Basic Argument tells us that there is simply no "work" left to be done by the natural laws when it comes to undermining free will.[15] At any point in time *t* at which someone exists, we can ask: "Is *t* the

---

[15] When G. Strawson makes claims such as "According to the Basic Argument, *it makes no difference* whether determinism is true or false" (2008: 289; my emphasis), it is not entirely clear whether he means that *both* sorts of laws pose a threat, *neither* do, or some alternative. One might think that Strawson means by this that his Basic Argument is equivalent to what is often called the "Standard Argument," an argument for impossibilism that roughly consists of the conjunction of the Consequence Argument and the *Mind* Argument (for a description of the *Mind* Argument, see van Inwagen 1983:126–152). However, I deny that the Basic Argument is equivalent to the Standard Argument. As noted above (fn. 12), it is unclear whether the Consequence Argument is best understood as an argument for incompossibilism or incompatibilism. However, the *Mind* Argument does identify *indeterministic causation* as a positive threat to free will. As such, the conclusions of the Basic Argument and the Standard Argument are not equivalent. Whether the explanatory conclusions of these two arguments are consistent seems to depend on the modal semantics being assumed. Assuming impossible worlds semantics, one might meaningfully claim that no possible being acts freely because no possible being satisfies the "starting point" sourcehood condition—but *if, counterpossibly, someone did satisfy it,* indeterministic causation would (if properly situated in the causal chain leading to action) undermine that person's freedom and moral responsibility. In that case, the conclusions of the two arguments are consistent. However, assuming standard possible worlds semantics, it seems that the conclusion of the Basic Argument and the conclusion of the Standard Argument are contrary claims: if the explanation of the impossibility of free will defended by the Basic Argument is correct, then there is no possible world at which indeterministic causation (of any sort) undermines free will, in which case the explanatory

starting point that constitutes an act of ultimate (ex nihilo) self-origination?" and the answer in each case must be "no." The answer will be "no" *irrespective* of whether someone S is living in a universe with deterministic laws or indeterministic laws. The answer will be "no" *irrespective* of whether S is an eternal being, a being with an infinite past, a being with no "remote past," or whether the being lives in a universe which has no past at all (prior to *t*).[16] As such, the Basic Argument is an argument for impossibilism, and *a fortiori* for incompossibilism, but the Basic Argument is not an argument for the incompatibilist view that deterministic laws (either alone or in combination something else) *undermine* free will. Indeed, if the Basic Argument is sound, it seems to follow that incompatibilism is *false*.

In rejoinder, some readers might argue that a philosopher could consistently accept both the soundness of the Basic Argument and the truth of incompatibilism, even if the Basic Argument is not an argument for incompatibilism. It is obvious that the proponent of the Basic Argument can and must endorse *incompossibilism*, a view which (as noted above) is often *called* "incompatibilism." However, it is far from obvious that a philosopher can coherently hold that deterministic laws *undermine* free will even though it is metaphysically impossible for someone to lack free will *in virtue* of being subject to deterministic laws. As such, something must be said in favor of accepting the view that the soundness of the Basic Argument is consistent with the truth of incompatibilism.

In defense of the consistency of such a view, a philosopher might forward an argument along these lines:

> Say that God exists and, for reasons having to do with his (necessary) nature and (necessary) value facts, God (necessarily) does not want to and, so, does not create any universe that does not have at least one free agent living in it. Let us say also that (necessarily) God knows that if he were to create a universe with deterministic laws, then any beings he were to create in that universe who were subject to those laws would not be free *because* being subject to deterministic laws undermines free will. It follows that necessarily, God does not create a universe with deterministic laws. Here, then, we have a scenario in which it is impossible for anyone to lack free will because they are subject to deterministic laws, yet it seems that incompatibilism is true. Indeed, it seems to be precisely because God recognizes that incompatibilism is true that God makes sure that deterministic laws never "get the chance" to undermine anyone's free will! Thus, the incompatibilist may be right in claiming that deterministic laws *preclude* or *undermine* free action and moral responsibility even if it is not metaphysically possible for someone to be unfree in virtue of being subject to deterministic laws.

Footnote 15 continued
conclusion of the Standard Argument is false; on the other hand, if there is some possible world at which indetermnistic causation (of some sort) undermines free will, then the explanatory conclusion defended by the Basic Argument is false.

[16] A *remote past* is a time "before there were any human beings" (van Inwagen 1989: 224).

The argument above, let us call it the "Weak Preclusion Argument," is a bit sketchy. Still, it suggests a way of understanding "precludes" that is sufficiently modest that a philosopher could consistently endorse the conclusion of the Basic Argument and the view that deterministic laws *preclude* free will. This line of reasoning suggests a "counterpossible" characterization of incompatibilism: If (perhaps counterpossibly) deterministic laws were to obtain, then one's evolving in accord with these laws would undermine one's free will so long as (perhaps counterpossibly) nothing else preemptively undermines one's free will. Assuming that this counterpossible characterization of incompatibilism is coherent, a philosopher may endorse incompatibilism and accept that the Basic Argument is sound.

However, the cogency of the counterpossible reasoning used in the Weak Preclusion Argument is disputable. The proponent of the argument assumes that we can reason non-trivially from an impossible antecedent, but explaining how this can be done has proven difficult (cf. Berto 2013). Assuming standard possible world semantics, every conditional with an impossible antecedent is *trivially* true: a conditional is true when its antecedent is false; when the antecedent is false at every possible world, then the conditional is true at every possible world regardless of the truth-value of its consequent. As such, according to standard possible worlds semantics, if there is no possible world at which the claim "God creates a universe with deterministic laws" is true, then the counterpossible conditional claim "If God were to make a universe with deterministic natural laws, then the deterministic laws of that universe would undermine the free will of everyone subject to the laws" is true, but trivially so. For the same reasons, though, the claim "If God were to make a universe with deterministic natural laws, then the deterministic laws of that universe would *not* undermine the free will of anyone subject to the laws" is also trivially true. Proponents of possible worlds semantics are happy with this result, for it reflects the intuitively compelling thought that if an impossibility were to obtain, then anything would follow. The proponent of the Weak Preclusion Argument, though, must deny that the former claim is vacuously true and must hold that the latter claim is false, and neither of these assessments is supported by standard possible worlds semantics.

Moreover, assuming standard possible worlds semantics, the incompatibilist claim "Deterministic laws undermine free will" is true only if there is some possible world at which it is true that deterministic laws undermine free will. But if we respect the stipulation made in the Weak Preclusion Argument that it is metaphysically impossible for God to create a universe with deterministic laws, then there is no possible world at which it is true that God creates a universe with deterministic laws. And if there is no possible world at which it is true that deterministic laws obtain, then there is no possible world at which it is true that deterministic laws undermine someone's free will. In that case, the claim that deterministic laws undermine free will is necessarily false—and (since that is standard possible worlds semantics' final say on the matter) this means that incompatibilism is false. So, assuming standard possible world semantics, if the Basic Argument is sound, then incompatibilism is false.

Notably, nothing in the discussion above indicates that the Weak Preclusion Argument is in fact unsound. Rather, what we have seen is that the proponent of this argument is assuming some non-standard (e.g. impossible worlds) semantics and, so, the argument will be every bit as controversial as the non-standard semantics being assumed—and every extant non-standard semantics is highly controversial. So, while I share the optimism of those who believe that philosophers will find a way to defend counterpossible reasoning, it is far from uncontroversial that the incompatibilists can endorse the conclusion of the Basic Argument.

The upshot of this discussion of the Basic Argument and Weak Preclusion Argument is that incompossibilism does not entail incompatibilism and that, depending upon one's preferred modal semantics, some arguments for incompossibilism are also arguments against incompatibilism. This undermines the view, proposed at the beginning of this section, that any sound manipulation argument for incompossibilism must also be a sound argument for incompatibilism. As such, we cannot simply assume that ZAM-2 is a sound argument for incompatibilism, even if we accept that it is a sound argument for incompossibilism.

5 A new version of the Zygote Argument

In this essay, I have argued that the original Zygote Argument concludes to incompossibilism, but is not an argument for incompatibilism (as I characterize these views). Still, the suggestion that the Zygote Argument could be retooled as a persuasive argument *against* incompatibilism may seem implausible. According to Mele's zygote story, Diana's act of creating zygote Z *guarantees* that Ernie will perform a particular action A at a specific future time and place. At first glance, it seems that Diana is only able to guarantee Ernie's future performance of A in this way because (she knows that) the causal chain leading from her creative act to Ernie's A-ing will be perfectly deterministic (at least at the macro level). As such, the reader might reasonably wonder how it could be that the deterministic causal chain between Diana's act of creation and Ernie's A-ing is not at least *part* of what accounts for Ernie's lack of freedom and responsibility for A-ing. In other words, given the central role played by causal determination in the zygote story, one might think that we can and should *add* some origination-related feature into our account of what undermines Ernie's freedom, but we cannot *drop* determination.

However, a philosopher might coherently accept the truth of both her intuitive judgment that Ernie lacks free will and responsibility and the empirical claim that the deterministic causal laws in Mele's zygote story play a crucial role in *eliciting* that intuitive judgment, and yet reject the incompatibilist's metaphysical claim that causal determination (either alone or in conjunction with something else) *undermines* free will and moral responsibility. It may be that *thinking about* a creation story which features deterministic laws (rather than indeterministic laws) helps us to get a better grasp on the ramifications of the freedom-undermining way in which we all come into existence. Because the laws of nature in Ernie's universe are deterministic, we can neatly trace Ernie's A-ing back in time to the properties of his zygote, properties over which Ernie exerted no control. In this way,

deterministic laws help us to pinpoint the *genuine* threat to free will and moral responsibility: failing to be the "ultimate source" of one's actions through an act of ultimate (ex nihilo) self-creation, i.e. *constitutive luck*. In short, a philosopher might argue that the Zygote Argument is best understood as *persuasive* version of the Basic Argument.

Fully fleshing out a best-explanation argument in support of an anti-incompatibilist, constitutive-luck version of the Zygote Argument would take some doing. However, in the light of my characterization of the Basic Argument (Sect. 4), it seems that there is logical space for such an argument to be developed. Indeed, Mele points in the direction of a constitutive-luck explanation when he emphasizes the "way in which his [Ernie's] zygote was produced" and "the way Ernie's zygote comes to exist" in the premises of the Zygote Argument and classifies the Zygote Argument as a "original design" argument (e.g. 2008: 278). Moreover, I am not the first to suggest that the upshot of manipulation arguments like the Zygote Argument may be that freedom-undermining constitutive luck is metaphysically impossible to overcome, irrespective of the natural laws. For instance, although Neil Levy does not discuss the Zygote Argument in his recent book *Hard Luck* (2011), he does consider manipulation stories quite similar to Ernie's. In these manipulation stories, Levy argues, the victims lack freedom and responsibility due to constitutive luck, and their problem with constitutive luck is not bound up with the type of causal laws that obtain (2011: 86–89). Levy's work is especially noteworthy in this context because Levy does not explicitly endorse or draw upon the contentious Basic Argument to make his case for impossibilism and against incompatibilism.

It should now be clear that those who wish to use Mele's zygote story in a defense of incompatibilism carry a heavier burden than is commonly recognized. Not only must the incompatibilist supplement the original Zygote Argument with a best-explanation argument which pinpoints deterministic laws as playing a role in undermining Ernie's freedom and responsibility, but they must also show that Ernie's lack of freedom and responsibility *cannot* be explained solely in terms of constitutive luck.[17] Even among those who are persuaded by Ernie's story that incompossibilism is true, the constitutive-luck explanation may have a special allure. Incompossibilism may be true, but "disappointed compatibilists" can at least

---

[17] The interesting dialectic between *incompatibilist*-impossibilists and *non-incompatibilist*-impossibilists has not been given much attention in the contemporary literature—presumably because the distinction between incompatibilism and incompossibilism has not been widely recognized. Kadri Vihvelin is the most significant exception. Among other things, Vihvelin (e.g. 2008, 2013) argues that impossibilism and incompatibilism are logically inconsistent views. If Vihvelin is right, then a philosopher could not, for example, argue for impossibilism by using (1) the Zygote Argument to defend incompatibilism with respect to free will and deterministic laws and (2) a second argument–e.g. the *Mind* Argument (van Inwagen 1983:126–152)—to defend the incompatibility of free will and *indeterministic* laws. While I have argued elsewhere (Mickelson forthcoming) that Vihvelin's arguments for the purported inconsistency of incompatibilism and impossibilism fail, I am sympathetic to the nearby view that any compelling best-explanation argument for incompatibilism must assume (or argue for) the metaphysical possibility of free action. In order to get clear on the lesson of the Zygote Argument, it seems that philosophers will have to consider such taxonomical and dialectical issues in greater detail.

take solace in the fact that they got something right: deterministic laws *do not* pose a threat to free will.[18]

6 Conclusion

In this essay, I have argued that the original Zygote Argument is invalid, but in instructive ways. First, the effort to identify the formal problems with the original version of the Zygote Argument led to a useful distinction between incompatibilism and mere incompossibilism. Once drawn, this distinction helped us to see that Mele's most recent version of the Zygote Argument is valid but, as an argument for mere incompossibilism, does not pinpoint deterministic laws as a threat to anyone's free will or moral responsibility. Then, by considering what would be required to upgrade a manipulation argument for mere incompossibilism to a diagnostic argument for incompatibilism, it became apparent that *no* manipulation argument for incompatibilism can be developed without appeal to an argument to the best explanation. Finally, I confirmed that there is room to develop the Zygote Argument as an argument for impossibilism but against incompatibilism. It sum, it seems that pigeon-holing the Zygote Argument as an argument about the freedom-relevance of deterministic laws significantly obscured the dialectic surrounding this and similar manipulation arguments. I expect that future discussion of the relative merits of these competing diagnostic versions of the Zygote Argument will help us to better understand the structures, limits, and lessons of global manipulation arguments.

References

Berto, F. (2013). Impossible worlds. In E. N. Zalta (Ed.), *The stanford encyclopedia of philosophy* (Winter 2013 Edition). http://plato.stanford.edu/archives/win2013/entries/impossible-worlds/
Campbell, J. (2007). Free will and the necessity of the past. *Analysis, 67*, 105–111.
Cornman J. and Lehrer K. (1968). *Philosophical Problems and Arguments: An Introduction*. New York: MacMillan Company.
Demetriou, K. (see also Mickelson, K.) (2010). The Soft-Line Solution to Pereboom's Four-Case Argument. *Australasian Journal of Philosophy*, 88(4), 595–617.
Kapitan, T. (2000). Autonomy and manipulated freedom. *Philosophical Perspectives, 14*, 81–104.

---

[18] I borrow the notion of a "disappointed compatibilist" from Levy (2011: 2). Levy says that he is "disappointed" insofar as he was hoping that compossibilism (rather than incompossibilism) would turn out to be true, but a "compatibilist" insofar as he denies the incompatibilist's claim that deterministic laws preclude free will.

Kearns, S. (2012). Aborting the Zygote Argument. *Philosophical Studies, 160*(3), 379–389.

Lehrer, K. (1960). *Ifs, cans, and causes*. Dissertation, Brown University. Providence: ProQuest/UMI. (Publication No. AAT: 6205755.)

Levy, N. (2011). *Hard luck: How luck undermines free will and moral responsibility*. New York: Oxford University Press.

McKenna, M. (2010). Whose argumentative burden, which incompatibilist arguments?—Getting the dialectic right. *Australasian Journal of Philosophy, 88*(3), 429–443.

McKenna, M. (2012). Moral responsibility, manipulation arguments, and history: Assessing the resilience of nonhistorical compatibilism. *Journal of Ethics, 16*(2), 145–174.

Mele, A. (2006). *Free will and luck*. New York: Oxford University Press.

Mele, A. (2008). Manipulation, and moral responsibility. *Journal of Ethics, 12*(3), 263–286.

Mele, A. (2012). Manipulation, moral responsibility, and bullet biting. *Presentation at the workshop on manipulation arguments*, Central European University. http://humanproject.ceu.hu/events/2012-06-07/workshop-on-the-manipulation-argument

Mele, A. (2013). Manipulation, moral responsibility, and bullet biting. *Journal of Ethics, 17*(3), 167–184.

Mickelson, K. (see also Demetriou, K.) (forthcoming). A Critique of Vihvelin's Three-fold Classification. *Canadian Journal of Philosophy*.

Strawson, G. (1986). *Freedom and belief*. Oxford: Clarendon Press.

Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies*, *75*, 5–24. Reprinted in Watson 2003.

Strawson, G. (1998, 2011). *Free will*. In E. Craig (Ed.), *Routledge Encyclopedia of Philosophy*. London: Routledge. http://www.rep.routledge.com/article/V014SECT3

Strawson, G. (2002). The bounds of freedom. In *Kane 2002* (pp. 441–460).

Strawson, G. (2008). The impossibility of ultimate moral responsibility. In D. Pereboom (Ed.), *Free will second edition 2009* (pp. 289–306); originally from 2008 *Real Materialism and Other Essays*. Oxford: Oxford University Press.

van Inwagen, P. (1969). *An Essay on the Freedom of the Will*. Dissertation, The University of Rochester. Rochester: ProQuest/UMI. (Publication No. AAT: 6914798.)

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

van Inwagen, P. (1989). When is the will free? *Philosophical Perspectives, 3*, 399–422. Repr. in *Agents, causes, and events*, ed. T. O'Connor. Oxford: Oxford University Press, 1995. Page references are to this latter work.

van Inwagen, P. (2008). How to Think About the Problem of Free Will, *Journal of Ethics*, *12*(3/4), 327–341.

Vihvelin, K. (2008). Compatibilism, incompatibilism, and impossibilism. In John Hawthorne, Theodore Sider, & Dean Zimmerman (Eds.), *Contemporary debates in metaphysics* (pp. 303–318). Malden, MA: Blackwell Publishing.

Vihvelin, K. (2013). *Causes, laws, and free will: Why determinism doesn't matter*. New York: Oxford University Press.