



The Zygote Argument Is Still Invalid: So What?

Kristin M. Mickelson^{1,2}

Received: 3 March 2020 / Revised: 26 June 2020 / Accepted: 30 June 2020 /
Published online: 3 October 2020
© The Author(s) 2020

Abstract

In “The Zygote Argument is Invalid: Now What?”, Kristin Mickelson argues that Alfred Mele’s original Zygote Argument is invalid: its two premises tell us merely that the truth of determinism is (perhaps spuriously) correlated with the absence of free human agents, but the argument nonetheless concludes with a specific explanation for that correlation, namely that deterministic laws (of the sort described by determinism) preclude—rule out, destroy, undermine, make impossible, rob us of—free will. In a recent essay, Gabriel De Marco grants that the original Zygote Argument is invalid for the reasons that Mickelson has identified, and claims that he has developed two new solutions to her invalidity objection. In this essay, I argue that both of his proposed solutions are nonstarters, the first fails as a “rescue” because it simply restates an extant solution in new jargon and the second fails because it consists in another invalid variant of the original Zygote Argument.

Keywords Free will · Determinism · Manipulation arguments · Best-explanation argument

1 Introduction

In “The Zygote Argument is Invalid: Now What?” (2015b), I argued that Alfred Mele’s original Zygote Argument (Mele 2006, 2008, 2012) is invalid: its two premises tell us merely that the truth of determinism is (perhaps spuriously) *correlated* with the absence of free human agents, but the argument nonetheless concludes with a specific *explanation* for that correlation, namely that deterministic laws (of the sort described by determinism) preclude—rule out, destroy, make impossible, conflict with, rob people of—free will. As such, the original Zygote Argument is deductively invalid (in classical logic), and to infer the Zygote Argument’s explanatory conclusion from its two

✉ Kristin M. Mickelson
kristin.mickelson.42@gmail.com

¹ University of Minnesota, Morris, 600 E 4th St, Morris, MN 56267, USA

² University of Gothenburg, Gothenburg, Sweden

premises is to commit the *cum hoc, ergo propter hoc* (with this, therefore because of/on account of this) fallacy.

In a recent essay, Gabriel De Marco (2016) grants that the original Zygote Argument is invalid for the reasons that I identified, and claims that he has developed two new solutions to my objection. In this essay, I argue that both of his proposed solutions fail. The first fails as a “rescue” because it merely repeats (in slightly different jargon) a candidate solution that I proposed in Mickelson 2015b. The second solution fails because it consists in another invalid variant of the original Zygote Argument.

This essay has four main sections. I begin (§2) with a brief overview of the traditional problem of free will and determinism, the problem that the Zygote Argument purports to solve. I then address De Marco’s first reply (§3) and second reply (§4) to my invalidity objection in turn. In closing (§5), I identify the role that covert best-explanation reasoning plays in De Marco’s second solution and offer a few reflections on the current state of the free-will debate.

2 The Problem of Free Will and Determinism

The traditional problem of free will and determinism has two basic components, the *correlation problem* and the *explanation problem* (Mickelson 2019a, 2019b, forthcoming). The correlation problem is the familiar challenge of settling the answer to this simple question: Is it metaphysically possible for a normal human to exercise free will when determinism is true? Free-will theorists have been deeply divided by this question for millennia, with many insisting that the answer is “yes” while others insist that the answer is “no.” Since a “no” answer is a commitment to the view that it is impossible for free human agents to exist in a world with deterministic laws (of the sort described by determinism), i.e. to the view that these two phenomena are *impossible*, let us refer to the “no” answer as an *impossibility solution* to the correlation problem. By contrast, a “yes” answer will be a *compossibility solution*.

Despite superficial appearances, the (in)compossibility debate is not simply a dispute about the modal distribution of free human agents. Rather, this debate is philosophically interesting because it is a debate about the nature of free will itself: anyone who finds the impossibility solution intuitive has thereby discovered that, by their intuitive lights, any successful account of free will must include (at least) one necessary condition which is impossible for a human to satisfy when determinism is true. Those who find the compossibility solution satisfying are, thereby, denying that there is any such desideratum on a successful account of free will. As such, the (in)compossibility dispute can be understood both as fundamental dispute about the nature of free will and a dispute about the standards by which a candidate account of free will should be judged.

Although an impossibility solution to the correlation problem is interesting and dialectically significant in its own right, it also raises a new and challenging *explanation problem*. This explanation problem consists in two closely related challenges:

The Unmet Condition Challenge (E1): Identify the necessary condition C on free will which humans cannot satisfy when determinism is true.

The Condition-Underminer Challenge (E2): Identify what keeps normal humans from satisfying condition C when determinism is true.

In order to solve (E2), one must identify the precise features/factors of normal humans and/or their environment which make it impossible for a person to satisfy the necessary condition named in the solution to (E1).

Candidate solutions to (E1) are usually divided into three main categories: *could-have-done-otherwise* conditions (a.k.a. *leeway* conditions), *source* conditions, and *hybrid* conditions (according to which an ability to do otherwise is what allows one to satisfy the source condition on free will). Since it is safe to assume these types of solutions to (E1) are already familiar to the target audience of this essay and they play no role in the present discussion, I will not review them here.¹

Within the mainstream free-will literature, philosophers do not have any standardized categories for candidate solutions to (E2). However, we can remedy this by importing into our discussion the three basic categories of “luck” introduced by Thomas Nagel in his early work on the paradox of moral luck.² Following Nagel, I will use the phrase “due to luck” as shorthand for “due to factors beyond one’s control.”³ The three standard categories of luck (i.e. factors beyond one’s control) are: *causal luck*, *circumstantial luck*, and *constitutive luck*. A person is subject to *causal luck*, as it is to be understood here, when the factors which account for the diachronic (state-by-state and moment-by-moment) evolution of the physical world—causal relations, causal laws, laws of nature, or the like—are beyond one’s control. Non-causal/actor-extrinsic *circumstantial luck* obtains when one’s action is due, at least in part, to circumstances (events, states of affairs, etc.) external to the actor that are beyond that actor’s control. *Constitutive luck* obtains when one’s actions are due, at least in part, to constitutive properties of the actor that are beyond the actor’s control.

We can easily categorize the most familiar types of solutions to (E2) based on which of these factors beyond one’s control is doing the *work* of keeping people from acting freely. A strict *causal luck solution* to (E2) proposes that it is impossible for humans to act freely when determinism is true *because* humans are subject to deterministic causal relations (deterministic causal laws, deterministic laws of nature, or the like). Peter van Inwagen’s Consequence Argument (van Inwagen 1983) is generally considered to be an argument for a causal-luck solution to (E2).⁴ By contrast, a strict *constitutive luck solution* to (E2) proposes that what keeps people from satisfying the condition named in (E1) is *constitutive luck*, i.e. the barrier to free agency is that the properties which constitute the agent (at least in key mental respects) are beyond the agent’s control. For example, Galen Strawson uses the Basic Argument to defend a strict constitutive-luck solution to (E2), claiming that that is constitutive luck alone which makes people

¹ For further discussion of these conditions, see McKenna and Pereboom 2016; for discussion of their important dialectical role in solving the explanation problem of free will and determinism, see Mickelson 2019a, 2019b, and “Hard Times for Hard Incompatibilism” (manuscript).

² The *paradox of moral luck* was introduced in a pair of essays by Bernard Williams and Thomas Nagel (Williams and Nagel 1976). For an accessible overview of the problem of moral luck see Hartman 2017, and for a discussion of the theoretical connections between the problem of moral luck and the problem of free will and determinism, see Mickelson 2019b.

³ In other words, I will assume that the so-called “lack of control” account of luck captures the notion of luck which is central to the moral-luck debate (cf. Anderson 2019, Hartman 2017: 23–31, Statman 2019).

⁴ However, the conclusion of the Consequence Argument is a growing matter of dispute in the wake of Joseph Campbell’s “No Past Objection” to the Consequence Argument (Campbell 2007).

unfree—from which it follows that the causal-luck solution to (E2) is false (cf. Strawson 1994; Mickelson 2015b, 2019b). A strict *circumstantial luck solution* to (E2) says that it is impossible to satisfy the condition on free will named in (E1) when determinism is true *because* people lack control over certain non-causal, agent-extrinsic events which obtained prior to their actions.⁵ Usually, though, philosophers who worry about circumstantial luck do not think that it is a lack of control over past circumstances *alone* that makes a person unfree, but rather combine circumstantial luck with some other type of luck to create a *hybrid solution* to (E2). Some philosophers have suggested that a circumstantial/constitutive luck solution could be correct (cf. Latus 2001, Nagel 1986: 113–14). Others have suggested a causal/circumstantial luck solution according to which it is the combination of our lack of control over the laws and circumstances in the remote past which makes us unfree now (cf. Bailey 2012: 361).

Our review of extant solutions to (E2) upsets the common assumption, popular in both the free-will literature and moral-luck literature, that broadly (i.e. strict and hybrid) causal luck solutions to (E2) are the only sort available. When it comes to the supposed impossibility of free will and deterministic laws, there is an in-house debate among impossibilists about whether this impossibility is spurious (as the constitutive luck solution proposes) or not (as a broadly causal luck solution proposes). This means that there is nothing peculiar to the dialectical structure of the problem of free will and determinism which allows one to safely infer a causal luck solution to (E2) from an impossibility solution to the correlation problem; to infer the former from the latter amounts to fallacious *cum hoc, ergo propter hoc* reasoning.

The reader will note that the terms ‘incompatibilism,’ ‘compatibilism,’ ‘compatibility,’ and ‘incompatibility’ have not been used in my characterization of the traditional problem of free will and determinism. This is because there is currently no consensus among philosophers about the meaning of these terms, so anyone who uses these terms must either assign a controversial definition to them or define them in such imprecise ways that the meaning of the terms are unclear. When pressed, most free-will specialists agree that the term ‘incompatibilism’ has traditionally picked out a broadly causal luck solution to (E2) of the explanation problem (cf. Pereboom 2001, 2014; Vihvelin 2008, 2013, 2018; McKenna 2010: 432; Levy 2011: 1; McKenna and Pereboom 2016: 151; Sartorio 2016: 147). This is consistent with the etymology of the term ‘incompatibilism’: this bit of jargon was (by all appearances) coined by Keith Lehrer to pick out roughly the thesis that it is *owing to determinism* that the free-will thesis is false when determinism is true.⁶ By contrast, Mele contends that he is “following standard practice” when he uses this term to name nothing beyond a mere impossibility solution to the correlation problem. He also proclaims (without

⁵ Carolina Sartorio (2015) comes close to proposing a strict circumstantial-luck solution in her discussion of backward-moving time travelers insofar as she suggests that their causal control over the past might allow them to be free even though we (as non-time-travelers) are not.

⁶ Lehrer has confirmed (in conversation and correspondence) that incompatibilism (as he understands it) is an explanatory view according to which the truth of determinism is *relevant* to the falsity of the free-will thesis in worlds at which determinism is true. Among other things, this means that Lehrer’s incompatibilism is not fully captured by the inconsistency thesis *that necessarily, determinism materially implies the negation of the free-will thesis*, nor is it captured by the impossibility thesis *that free will is impossible (for unspecified reasons) with the sort of causal laws described by determinism*. This amends the etymological claim made in Mickelson 2015b, but lends support to the explanatory characterization of ‘incompatibilism’ that I endorse.

argument or evidence) that philosophers who define the term otherwise are guilty of defining the term in a “nonstandard” or “nontraditional” way (Mele 2017: 6, n. 4, and 2019: 3, n.1).⁷ In order to sidestep this dispute about who is “correctly” defining bits of technical jargon, I will not use the term ‘incompatibilism’ or its cognates to make any of the central points in this essay; I will mention them only when doing so contributes to my assessment of De Marco’s proposed solutions to the invalidity objection.

3 De Marco’s Terms and “Initial Solution”

Mele has presented several statements of the Zygote Argument since it was first introduced, and it will be useful to adopt an intuitive way of tracking them. With that aim in mind, I will hereafter refer to statements of the Zygote Argument by the year they first appeared in print, starting with “ZA-2006” to label the very first statement of the Zygote Argument (Mele 2006).⁸ My invalidity objection targets a later statement of the Zygote Argument that Mele put forward in response to criticisms of ZA-2006, hereafter “ZA-2012”⁹:

ZA-2012

P1. Ernie is not [free or] morally responsible for anything he does.

P2. 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.

C. So determinism precludes [free will and] moral responsibility – at least for human beings who develop from normal human zygotes. (Mele 2012: 15)¹⁰

⁷ Pace Mele, has not consistently used this definition of ‘incompatibilism.’ First, Mele uses the standard ambiguous definitions of ‘compatibilism’ and ‘incompatibilism’ (such that *compatibilism* as the view that free will is *compatible* with determinism and *incompatibilism* as the view that free will is *incompatible*, where the nature of the compatibility and incompatibility relations is left undefined) at the outset of the monograph in which the Zygote Argument was introduced (Mele 2006: 1; see also Mele 1995: 9). Second, Mele uses ‘incompatibilism’ as Pereboom does (i.e. in the explanatory, etymological sense) when he claims that Pereboom’s argument “fails as an argument for incompatibilism” is untenable when ‘incompatibilism’ picks out a mere impossibility claim (cf. Mele 2005, 2006: 144). The ways in which Mele accepts/promotes different definitions of the term ‘incompatibilism’ in different dialectical contexts makes him vulnerable to a charge of fallacious motte-and-baileying, a charge I develop elsewhere (“Motte-and-Bailey Incompatibilism,” manuscript).

⁸ Mele used the shorthand “ZA” as the specific name for ZA-2006 and has since introduced labels such as “ZAM” and “ZAM2” for new variant of the argument. I prefer a labelling strategy that makes the timeline of the changes evident.

⁹ Mele calls ZA-2012 by the label “ZAM” and uses the same label for ZA-2013; in Mickelson 2015b, “ZAM-1” is the label for ZA-2012, and “ZAM-2” the label given to ZA-2013. It is not clear to me that ZA-2012 is the same argument as ZA-2006, but I will grant Mele’s claim that ZA-2012 is just a clearer statement of ZA-2006, from which it follows that any substantive criticism of ZA-2012 applies equally to ZA-2006.

¹⁰ It is generally accepted that the manipulation victim in any properly designed manipulation will satisfy all epistemic conditions on moral responsibility, such that the victim lacks moral responsibility in virtue of failing the control condition (i.e. the free-will condition) on moral responsibility (cf. Mickelson 2017). For the sake of simplicity, I will assume that ZA-2012’s conclusion, like that of ZA-2006, tells us that determinism precludes—undermines, conflicts with, makes impossible—free action and, so, moral responsibility.

ZA-2012 is invalid: its premises entail merely that it is metaphysically impossible for a normal human born from a human zygote to act freely and morally responsibly when determinism is true, but its conclusion says that it is *owing to determinism* that people lack free will and moral responsibility when determinism is true.¹¹

So understood, we can say that ZA-2012 is invalid because its premises support only (an idiosyncratically restricted variant of) the impossibility solution to the correlation problem, but its conclusion forwards a controversial causal luck solution to (E2) of the explanation problem.

Having shown that ZA-2012 is fallacious, I outlined two ways to improve the original Zygote Argument, saying:

[T]here 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 impossibilism [i.e. the thesis that deterministic laws and free action are impossible]. 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. (Mickelson 2015b: 2912; my emphasis)

Here, I propose two distinct strategies for responding to the invalidity objection. First, one may hold the two premises of ZA-2012 fixed and change its solution to the impossibility thesis entailed by those premises. For clarity, let us call any argument which results from implementing this “simplest repair” strategy a *premise-preserving solution* to the invalidity objection. Second, one may keep the premises of ZA-2012 fixed and update its premises so that they support this explanatory conclusion in order to create a variant of the Zygote Argument which is invulnerable to my invalidity objection. Let us call any argument which results from this strategy a *conclusion-preserving solution* to the invalidity objection. Notably, I argue that the only dialectically felicitous way to develop a *conclusion-preserving solution* to my invalidity objection is to supplement ZA-2012’s original premises with a best-explanation argument (Mickelson 2015b: 2912).

From the outset, I worried that philosophers might be tempted to brush off my complaints about the invalidity of ZA-2012 given how easy it is to develop a new, valid variant of ZA-2012 using the premise-preserving strategy (Mickelson 2015b: 2917). My proposed premise-preserving variant of the Zygote Argument, hereafter “Simple Repair,” may be summarized as follows:

Simple Repair

P1. Ernie does not freely A and is not morally responsible for A-ing.

P2. Concerning free action and moral responsibility, there is no significant difference between Ernie’s A-ing and any candidate for a free and morally responsible action in a deterministic universe.

¹¹ In a footnote, De Marco says that “some” might dismiss my claims that OZA is invalid by simply rejecting my interpretation of ‘precludes’ (De Marco 2016: 1623, fn.4). Perhaps so, but such a claim is neither an argument against the accuracy of my interpretation nor a defense of OZA’s validity. Speaking against such a dismissal is that my interpretation is consistent with both the lexical definition of ‘precludes’ and Mele’s use of the term (e.g. Mele 2017: 146, 148, 177; Mele 2013: 18; Mele 1995: 18).

C. So, no normal human candidate for a free and morally responsible action in a deterministic universe is a free action nor an action for which the actor is morally responsible. (Mickelson 2015b: 2917-2919)

In response to the invalidity objection and my overview of the two response strategies,¹² Mele adopted the following premise-preserving variant of ZA-2012, hereafter “ZA-2013”:

ZA-2013

P1. Ernie is [not a free agent and is] not morally responsible for anything.

P2. Concerning [the 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.

C. So in no possible deterministic world in which a human being develops from a normal human zygote is that human being [free or] morally responsible for anything he or she does (Mele 2013: 176).

Mele has recently reaffirmed his commitment to the premise-preserving strategy in his restatement of ZA-2013 as a *modus ponens* argument that I will call “ZA-2019”:

ZA-2019

P1. Ernie is not [a free agent or] morally responsible for anything he does.

P2. If Ernie is not [a free agent or] morally responsible for anything he does, no human being who develops from a normal human zygote in a deterministic world is morally responsible for anything he or she does.

C. So in no possible deterministic world in which a human being develops from a normal human zygote is that human being [a free agent or] morally responsible for anything he or she does. (Mele 2019: 120-121)

ZA-2019 is, uncontroversially, a valid argument.

Mele presents ZA-2019 in response to the worry that ZA-2013 is invalid.¹³ Mele’s attention to this invalidity objection is somewhat puzzling. Mele offers no evidence that any professional philosopher has had this sort of worry about the validity of ZA-2013, nor does he mention that multiple free-will specialists (e.g. Mickelson 2015b: 2917, n.6; Sartorio 2016: 163–164) had already addressed this “worry” by pointing out that arguments such as ZA-2013 may be represented as *modus ponens* arguments. Moreover, Mele’s attention to this relatively trivial invalidity objection to ZA-2013 might confuse the average reader, for he does not distinguish it from my invalidity objection

¹² I first presented the invalidity objection and my two response strategies in 2012 while serving as a commentator on Mele 2012 at CEU’s *Workshop on The Manipulation Argument* (<https://philosophy.ceu.edu/events/2012-06-07/workshop-manipulation-argument>). Mele’s premise-preserving solution to the objection appears in Mele 2013 (the published version of Mele 2012).

¹³ In Mickelson 2015b, ZA-2013 has the label “ZAM-2”. Since Mele now uses the label “ZAM2” for ZA-2019, I do not follow the Mickelson 2015b labelling system here.

to ZA-2012—which is noteworthy because restating ZA-2013 as ZA-2019 does not address *my* invalidity objection.¹⁴

To stave off needless confusion, let us draw a clear line between my invalidity objection to ZA-2012 and Mele’s trivial invalidity worries about ZA-2013. Hereafter, let the latter type of invalidity objection be known as *the logic-text objection*, since it amounts to little more than the (arguably misguided¹⁵) complaint that Mele should have stated ZA-2013 as a logic-text proof in classical deductive logic from the outset (rather than leaving this simple task as an exercise for his readers). By contrast, when an argument is invalid because it fails in its aim to close the “explanatory gap” between an impossibility solution to the correlation problem and a specific solution to (E2) of the explanation problem, I will call it an *explanatory gap objection*.¹⁶

While I grant that ZA-2013 and ZA-2019 are valid arguments, I do not think that Mele has, by adopting these new variants of ZA-2012/ZA-2006, adequately responded to the explanatory gap objection. Even if sound, ZA-2013 and ZA-2019 do nothing to adjudicate between the rival candidate *explanations* for the impossibility of deterministic laws and free human agents. I have motivated this point by demonstrating that ZA-2013 may be used as a foundation for a more sweeping generalization argument which concludes that free will is impossible due entirely to constitutive luck, from which it follows that a strict constitutive luck solution to (E2) is *true*, that the causal luck solution to (E2) promoted by ZA-2012 is *false*, and that the conclusion of ZA-2013 expresses a true but metaphysically arbitrary claim (Mickelson 2015b; see also Mickelson 2017, 2019a, and 2019b). To be clear, I have never proposed that ZA-2013 is invalid, unimportant, or otherwise uninteresting; I have simply pointed out the generally overlooked fact that the premises of ZA-2012 fall short of delivering a broadly causal luck solution, or any solution whatever, to (E2).¹⁷ This observation does not suggest that the Zygote Argument is unworthy of attention, but it does suggest that the underlying logical structure and upshot of manipulation arguments are poorly understood and, so, are worthy of considerably more attention than they have been given.

De Marco (2016) enters the debate at this point, claiming that he has devised two ways to “rescue” the Zygote Argument from my explanatory gap objection:

¹⁴ This is not a merely theoretical concern: I have refereed a manuscript which mistakenly stated that my invalidity objection targeted arguments with the logical form of ZA-2013.

¹⁵ The underlying logical structure of standard manipulation (a.k.a. multiple-case) arguments is not deductive, so representing them as *modus ponens* arguments may generate needless confusion about the overall logical structure and conclusion of manipulation arguments (cf. Mickelson 2015b, 2019a, “Hard Times for Hard Incompatibilism,” manuscript; see also Pereboom 2014: 79, n.3; van Inwagen 1983: 131).

¹⁶ Notably, Mele’s critique of Pereboom’s Four-Case Argument also qualifies as an explanatory gap objection, for Mele claims that the argument “fails as an argument for incompatibilism” on the grounds that Pereboom fails to defend his preferred diagnosis for the impossibility of free will and determinism (Mele 2005, 2006, 2008; Mickelson “Motte-and-Bailey Incompatibilism,” manuscript).

¹⁷ Although Mele has made no comments on the matter in print, he has publicly stated that he is uninterested in what accounts for/explains the lack of free agency when determinism is true (i.e. he is uninterested in which solution to (E2) of the explanation problem is correct), and this is why he is unconcerned by the modesty of ZA-2013’s conclusion. De Marco echoes Mele’s position when he claims “Whether these [manipulation] arguments prove incompatibilism or impossibilism is not very important for the debate” (De Marco 2016: 1624).

According to Mickelson, *the only way to make such arguments valid* is to supplement them with an argument that is an inference to the best explanation. In this paper, I argue that there are *two other ways* in which the proponent of such manipulation arguments can modify their argument, neither of which requires an inference to the best explanation. (De Marco 2016, abstract; my emphasis)

Pace De Marco, this summary of my position is flawed. It is not true that I say that the *only way* to change the invalid ZA-2012 into a valid argument is to “supplement them with an argument that is an inference to the best explanation.” As already discussed, I clearly outlined *two ways* to repair ZA-2012 in the face of the explanatory gap objection in Mickelson 2015b, namely the premise-preserving strategy and the conclusion-preserving strategy. I did claim that the only way to a *conclusion-preserving solution* to the explanatory gap objection to ZA-2012 is to add a best-explanation argument, but I also explicitly describe a premise-preserving solution that does not involve best-explanation reasoning of any kind (namely, “Simple Repair” above).

De Marco’s mischaracterization of my position is important when we look at his “two other ways” to repair manipulation arguments which are vulnerable to the explanatory gap objection.¹⁸ De Marco summarizes his “initial solution” to the explanatory gap objection as follows:

For those manipulation arguments that are invalid for the reasons that Mickelson gives [i.e., those that are subject to an explanatory gap objection], as well as for the Zygote Argument, an easy fix would be to change their conclusion to the claim that *compatibilism is false*. (De Marco 2016:1624, my emphasis)

Since De Marco stipulates his own definition for the term ‘compatibilism’ in the course of his 2016 essay, it is important to note that he uses the phrase “compatibilism is false” to pick out the following impossibility thesis:

there is no possible universe in which deterministic laws obtain and someone who is subject to the laws performs a free action (De Marco 2016: 1624).¹⁹

So, piecing things together, De Marco’s “initial solution” to the explanatory gap objection consists in the following premise-preserving variant of ZA-2012, call it “Easy Fix”:

Easy Fix

P1. Ernie does not act freely or responsibly.

P2. With regard to free action and moral responsibility, there is no significant difference between Ernie and a standard agent (i.e. someone who is, like Ernie, subject to the laws of nature) in a deterministic universe.

¹⁸ A person could fail to correctly identify my two replies but still successfully present two novel solutions (i.e. solutions not already in print), but this is not what De Marco has done.

¹⁹ De Marco also states, “I suggest that we take ‘compatibilism’ and ‘compossibilism’ to refer to the same thesis: the negation of impossibilism” (De Marco 2016: 1624).

C. So, there is no possible universe in which deterministic laws obtain and a standard agent performs a free action.²⁰

Easy Fix is a valid argument, so it may appear that De Marco's initial solution to the explanatory gap objection is a success.

Stepping back, however, grave problems with De Marco's first solution become apparent. The reader will note that the terms 'compatibility,' 'incompatibility,' 'impossibility,' 'compatibilism,' 'incompatibilism,' and 'impossibilism' do not appear in any of the statements of the Zygote Argument we have considered so far, i.e. ZA-2006, ZA-2012, ZA-2013, ZA-2019, Simple Repair, and Easy Fix. As such, how one defines *these terms* is completely irrelevant to the validity of *these arguments*. By extension, how one defines the aforementioned terms is also irrelevant when assessing which of these manipulation arguments is vulnerable to an explanatory gap objection. It is irrelevant, then, that De Marco refers to the impossibility thesis stated in the conclusion of Easy Fix by using the phrase "compatibilism is false" while I prefer the phrase "impossibilism is true" (for reasons that I provide in Mickelson 2015a, 2017, 2019a, 2019b, and forthcoming). This means that we can safely set aside De Marco's disapproval of my preferred labelling system²¹ and refocus on De Marco's claim that Easy Fix constitutes a new solution to my explanatory gap objection to ZA-2012.

Once we focus on the content of De Marco's Easy Fix, the differences between it and my Simple Repair disappear: the two arguments are constituted by the same three propositions presented in the same logical form. Simply put, Easy Fix and Simple Repair *are the same argument*.²² It follows that De Marco's first solution to the explanatory gap objection is a nonstarter.²³

4 De Marco's Second Solution

In De Marco's second response to the explanatory gap objection, he takes up a conclusion-preserving strategy. De Marco agrees that the original Zygote Argument's explanatory conclusion *may* be defended by adding a best-explanation argument to the premises of the original Zygote Argument (De Marco 2016: 1627), but he rejects my proposal that best-explanation reasoning is *required* to support its explanatory conclusion (i.e. that such reasoning is required to felicitously fill the explanatory gap

²⁰ This argument is based on De Marco's characterization of the "standard form" of manipulation arguments (De Marco 2016: 1623).

²¹ De Marco's criticisms of my proposed taxonomy of free-will views does not take into account the literature devoted to the difficult project of identifying the best taxonomy of free-will views (especially discussion of Kadri Vihvelin's "three-fold classification" found in Vihvelin 2008 and 2013, McKenna 2010, and Mickelson 2015a; see also Mickelson forthcoming), nor does De Marco show sensitivity to the extant criticisms of the type of impoverished taxonomy of free-will views that he adopts. As such, De Marco does not meet his argumentative burden in proposing that his use of jargon is superior to mine.

²² Moreover, it is overdetermined that Easy Fix is not a novel reply to the explanatory gap objection: Easy Fix differs only slightly from ZA-2013, so even if (counterfactually) I had not beaten De Marco to the punch, Mele clearly did.

²³ This claim rests upon the assumption that one does not give a *new argument* simply by assigning a *new label* to the proposition expressed in the conclusion of an extant argument. My thanks to a blind referee (at another journal) for revealing that some philosophers reject this methodological assumption and, thereby, that I should make my commitment to this assumption explicit.

between an impossibility solution to the correlation problem and a specific solution to (E2) of the explanation problem) (De Marco 2016: 1625). Rather than identifying a flaw in the argument I make in support of this proposal (Mickelson 2015b: 2912), De Marco proposes a counterexample in the form of a new variant of the Zygote Argument.

De Marco claims that a conclusion-preserving variant of the original Zygote Argument can be created without using best-explanation reasoning by adding a third premise which asserts that *free-will possibilism* is true, i.e. that it is metaphysically possible for someone to act freely (De Marco 2016: 1625–1626). De Marco does not provide a simple statement of this new manipulation argument, but the argument he describes, call it “ZAP-2016,” may be summarized as follows:

ZAP-2016

P1. Ernie does not freely A and is not morally responsible for A-ing.

P2. Concerning free action and moral responsibility, there is no significant difference between Ernie’s A-ing and a standard agent (i.e. a normal human who is, among other things, subject to the laws of nature) in a deterministic universe.

P3. Possibilism: Possibly, a standard agent performs a free action for which he is morally responsible.

C. So, determinism precludes—rules out, makes impossible, destroys, prevents the exercise of—free will and moral responsibility of standard agents in worlds at which determinism is true, i.e. it is impossible for a standard agent to exercise free will and moral responsibility (at least in part) *because they are subject to deterministic laws of nature*.

ZAP-2016 is invalid. Like ZA-2012, there is an explanatory gap between the proposition that is entailed by the premises and the proposition that is expressed by the conclusion.

Since ZAP-2016’s invalidity is not immediately obvious to everyone, some commentary on its invalidity is called for. The reader will note that ZAP-2016’s premises entail that possibly, someone exercises free will in a universe *without* deterministic laws, but they do not tell us *why* no one acts freely in universes *with* deterministic laws. For all that is said in the premises of this argument, it may simply be a *brute fact* that no humans act freely when determinism is true. This means that although ZAP-2016 was specifically introduced to solve the explanatory gap objection ZA-2012, it nonetheless succumbs to it: the relevance-relation-affirming “because” clause in ZAP-2016’s conclusion does not follow from the argument’s three premises. ZAP-2016 is deductively invalid, and to infer its conclusion from its premises is to commit the *cum hoc, ergo propter hoc* fallacy.

Clearly there are norms and narratives in place which make it easy for free-will theorists to overlook—or at least think it is wise to ignore—the fallacious reasoning found in arguments such as ZA-2012 and ZAP-2016, so it will be useful to more fully expose the invalidity in ZAP-2016’s underlying logical form. In giving ZA-2016, De Marco is *de re* proposing that the conjunction of possibilism and the impossibility solution to the correlation problem guarantees that a broadly causal luck solution to (E2) is correct—and, by extension, that all rival solutions to (E2) are incorrect. As such, we may summarize De Marco’s commitments by restating ZAP-2016 as follows, hereafter “ZAP-2020”:

ZAP-2020

P1. The impossibility solution to the correlation problem is true (assuming that both premises of the original Zygote Argument are true).

P2. Possibilism, the view that free will is metaphysically possible, is true.

C. A broadly causal luck solution to (E2) of the explanation problem is correct; non-causal solutions, such as the strict constitutive luck solution, are incorrect.

De Marco implies that there is no value in calling the substantively difference propositions expressed in P1 and C of ZAP-2020 by different names (De Marco 2016: 1623–1624). This is surprising, since he grants that it is fallacious to infer C from P1 alone (he grants this when he grants that ZA-2012 is invalid). By contrast, I advise that the propositions expressed in C and P1 should be given distinct names on the grounds that free-will theorists otherwise tend to conflate these substantively different claims and overlook the fact that it is fallacious to infer C from P1. I also think that the term ‘incompatibilism’ is an apt name to pick out the impossibility thesis stated in P1, and introducing a new term for P1 allows us to continue using ‘incompatibilism’ (as it traditionally has been) to name roughly the explanatory thesis stated in C.

Using my preferred terminology and taxonomy of free-will views, we can give an even more incisive summary of the underlying structure of ZAP-2016 and ZA-2020, an argument I will call the “Core Argument”²⁴:

The Core Argument²⁵

P1. Possibilism.²⁶

P2. Incompatibilism.²⁷

²⁴ In Mickelson 2015b, I consider views consisting in the same two-place relation but different relata. I used ‘incompatibilism*’ and ‘incompatibilism*’ for the views mentioned here in order to distinguish them from views which have no scope restriction on the agents of interest. Since I gloss over these scope restrictions on the free-will relata here (although I explain why they are critical to the free-will debate in Mickelson 2019a, 2019b, forthcoming), I will also drop the asterisk. See also note 26 below.

²⁵ ZAP-2020 seems to be a fair interpretation of the new Zygote Argument that De Marco has in mind. In the course of his reply, De Marco presents the following three theses:

D. [Incompatibilism]: Necessarily, anyone who is subject to deterministic laws is unfree.

E. [Incompatibilism]: D is true because these agents are subject to deterministic laws.

N. [Impossibilism]: Necessarily, no agent acts freely or responsibly.

De Marco contends that D entails E on the assumption that *N is false* (De Marco 2016: 1625), but this is not the case.

²⁶ Here, ‘possibilism’ names the view that it is metaphysically possible for a normal human to act freely.

²⁷ Here, ‘incompatibilism’ names the view that it is impossible for a normal human to act freely when determinism (roughly the thesis that, due to the facts of the past and the factors which govern the state-by-state evolution of the world—causal relations, causal laws, laws of nature, or the like—fix one unique and inevitable future) is true. Given my preferred definition of ‘determinism,’ there is no reason to include a “subject to the laws” restriction in this statement of incompatibilism, for it would be redundant: it follows from the definition of ‘determinism’ that *everyone* is subject to the laws on the assumption that determinism is true (cf. Mickelson 2019a, 2019b, and forthcoming). This unusual restriction was rhetorically useful in the narrow context of discussing ZA-2012 because, on the natural reading of Mele’s zygote story, Diana performs a miracle relative to the laws of nature when she creates Ernie’s zygote. If this is right, then Diana is not subject to the deterministic laws which obtain in Ernie’s universe, from which it follows that determinism (as defined above) is false. This is important, for some philosophers have suggested that a being who is not subject to the laws may act freely even though beings who are subject to the laws cannot (cf. Pereboom 2001, 2014; Stone 1998; Mickelson 2019a and “Hard Times for Hard Incompatibilism,” manuscript). My “subject to the laws” restriction was a charitable way to restrict the conclusion of my restatements of the Zygote Argument so that they did not suggest that (apparently) miracle-working beings like Diana lack free will, while sidestepping a discussion of Mele’s definition of the term ‘determinism’ and the meaning of the related phrase “deterministic laws.”

C. So, Incompatibilism.²⁸

The premises of the Core Argument entail that possibly, someone exercises free will in a universe *without* deterministic laws,²⁹ but they do not tell us why someone cannot act freely in universes *with* deterministic laws. The premises of the Core Argument do not tell us that deterministic causation or deterministic laws, *qua* being deterministic, pose a direct threat to free will. Nor do the premises of the Core Argument tell us that there is something special about the subset of metaphysically possible beings *who are subject to the laws* which make them unfree.³⁰ That is, the premises of the Core Argument do not indicate that the “subject to the laws” constraint on impossibilism (the view De Marco picks out with the phrase “compatibilism is false”) tracks a freedom-relevant feature of agents.³¹ In other words, the premises of the Core Argument do not tell us whether it is *relevant* to one’s ability to act freely that one is (or is not) subject to the laws, yet the conclusion clearly asserts that it is partly *because* people are subject to deterministic laws that they lack free will when determinism is true. In other words, there is still a sizable explanatory gap between the explanatory claim expressed in the Core Argument’s conclusion and the more modest non-explanatory conclusion entailed by the Core Argument’s premises. The Core Argument is subject to the explanatory gap objection—and, by extension, so is ZAP-2016 and ZAP-2020.

Having seen that ZAP-2016 is patently subject to the same type of explanatory gap objection as the original Zygote Argument, De Marco’s second attempt to rescue Mele’s argument from the explanatory-gap objection does not merely fall short of its aims; it falls short of qualifying as a viable attempt to address the problem. Moreover, since ZAP-2016 is invalid, this argument does not, *pace* De Marco, give us “reason to reject Mickelson’s thesis that such arguments need to be supplemented with an inference to the best explanation in order to be valid” (De Marco 2016: 1625).³² Indeed, De Marco provides the reader with no reason to question my argument for

²⁸ Here, ‘incompatibilism’ is used in roughly the traditional way, to name the view that determinism is incompatible with—precludes, undermines, destroys, makes impossible—human free will, from which it follows that it is impossible for a normal human to perform a free action when determinism is true (at least in part) *because* the person is subject to the type of causation and/or laws of nature described by determinism (as ‘determinism is defined in note 26 above). De Marco does not object to this use of the term ‘incompatibilism’; he simply sees no reason to “fuss” over the distinction I draw between incompatibilism (so characterized) and impossibilism (cf. De Marco 2016: 1622–1634).

²⁹ The following seems to be good reasoning:

1. Possibilism
2. Impossibilism
3. Some libertarian-friendly account of free will is correct and, possibly, someone satisfies it.

I discuss the role of libertarian-friendly accounts of free will in De Marco’s argument more in the next section.

³⁰ That is, there is something about the *deterministic* quality of the laws (causation, evolution, or the like) which keeps people from satisfying (at least) one of the necessary conditions on free will.

³¹ Whether one should think this scope restriction is arbitrary or not depends partly on one’s preferred account of free will and the definition of ‘determinism’ that one adopts (see note 27 above). Since De Marco provides no discussion of the dialectical significance of this scope restriction, I will not discuss its importance here.

³² Moreover, the pressing question is not fundamentally about deductive validity. The issue is better seen as whether the core reasoning for incompatibilism must be abductive/best-explanation reasoning, in which case any logic-text summary of the argument is just to give a deductive façade to an argument the quality of which turns on non-deductive reasoning. (See note 15 above).

the claim that closing the explanatory gap between an impossibility solution to the correlation problem and a specific candidate solution to (E2) of the explanation problem—at least in the context of manipulation arguments—requires a best-explanation argument.

5 Minimizing Vs. Closing the Explanatory Gap

Although De Marco’s conclusion-preserving solution technically fails, the reader might think that it still adequately addresses the explanatory gap objection: adding the assumption of possibilism to impossibilism does not *close* the explanatory gap between an impossibility solution to the correlation problem and a causal-luck solution to (E2) of the explanation problem, but perhaps it *minimizes* the gap to the point that it is no longer worrisome. After all, if possibilism is true and so is impossibilism, what other than deterministic laws could account for the lack of free human agents in deterministic universes? What *better explanation* could there be?

This question is worth raising because it exposes the problem with De Marco’s claim that he does not partake in best-explanation reasoning by adding the assumption of possibilism to the premises of ZA-2012:

Notice that nothing of what I have said above [in constructing ZAP-2016] amounts to an inference to the best explanation. Adding C [the assumption of possibilism] to the Zygote Argument [ZA-2012], or to manipulation arguments for incompatibilism [i.e. arguments for a broadly causal-luck solution to (E2) of the explanation problem], would not thereby make it a best explanation argument, nor would it need to be supplemented by one. (De Marco 2016: 1626)

Contrary to what De Marco’s commentary implies, there is more to a best-explanation argument than asserting a speculative explanatory claim. Among other things, one must build up a case for one’s preferred explanation by carefully considering viable alternative explanations and showing that one’s preferred explanation is *better* than its rivals.

Instead of taking up this difficult work (cf. Mickelson 2019a, 2019b), De Marco takes a dialectical shortcut: adding possibilism to the premises of ZA-2012 allows him to *reject out of hand* a swath of explanatory claims which imply that free will is impossible for reasons which have nothing whatever to do with the laws of nature. For instance, I have demonstrated that ZA-2013 may be used as the foundation for a manipulation which supports G. Strawson’s constitutive luck solution to (E2) of the explanation problem, from which it follows that *all broadly causal luck solutions are false* (Mickelson 2015b: 2923; see also my “master manipulation argumen” in Mickelson 2019b). To assume that possibilism is true, however, is to assume that G. Strawson’s preferred explanation for the impossibility free will—and, *a fortiori*, its spurious impossibility with *every possible phenomena* (from fluffy kittens to atom bombs and deterministic laws)—is false.

In the light of the discussion above, the dialectical role played by ZAP-2016’s possibilism-affirming premise is now clear: this premise rules out all extant candidate

solutions to (E2) which conflict with traditional causal luck solutions to (E2).³³ By adding possibilism to the premises of ZA-2012, De Marco clears away rivals to the broadly causal-luck explanation for the impossibility without having to engage with them. He thereby clears the way for a *best-explanation inference* to the conclusion that deterministic laws undermine human free will. *Pace* De Marco, then, adding possibilism to the premises of the original Zygote Argument to create an argument incompatibilism (i.e. a broadly causal luck solution to (E2) of the explanation problem) is to engage in a type of best-explanation reasoning after all.³⁴

Moreover, as the careful reader may have already noticed, adding possibilism to the premises of ZA-2012 (i.e. to the assumption that the impossibility solution to the correlation problem is correct) is, in effect, to assume that some *libertarian-friendly account of free will* is correct.³⁵ By “libertarian-friendly,” I simply mean those accounts of free will which include at least one necessary condition which *cannot* be satisfied by a normal human when determinism is true, but *can* be satisfied (along with all the other necessary conditions on free will) when determinism is false. De Marco gives us no reason to think that it is dialectically felicitous to assume that a libertarianism-friendly account of free will is correct in a manipulation argument for incompatibilism. That is, De Marco gives us no reason to suppose that it is dialectically felicitous to assume that some libertarian-friendly solution to (E1) of the explanation problem is correct in an argument which purports to show that an impossibility solution to the correlation problem is true, i.e. that all compossibility-friendly accounts of free will are false.

At precisely what point in the unfolding of a manipulation argument—if any—is it dialectically felicitous for a philosopher to appeal to their preferred libertarian-friendly account of free will to make their case for the causal luck solution to (E2) and against, for instance, a strict constitutive luck solution? Is a broadly causal luck explanation for the impossibility of free will and deterministic laws available only to those who are *already committed* to the truth of a specific libertarian-friendly view of free will?³⁶ Can defenders of *any* type of libertarian-friendly account of free will (e.g. both event-causal and agent-causal views) accept the premises and conclusion of the Core Argument, or only a subset? That is, can those who embrace a libertarian-friendly account of free will argue for a broadly causal luck solution to (E2) of the explanation problem without begging the question against both those philosophers who accept a compossibility solution to the correlation problem as well as those impossibilists who (like G. Strawson) reject libertarian-friendly

³³ For a brief survey of some of these candidate explanations, see Mickelson 2019a and “Hard Times for Hard Incompatibilism” (manuscript).

³⁴ There is good precedent for using phrases like “best explanation argument” and “best explanation reasoning” in this broad way. Pereboom (2001, 2014) and Mele (2013) say that the Four-case Argument includes a best-explanation argument even though Pereboom relies solely on a slippery-slope argument to *rule out* alternatives to his proposed solution to (E2) of the explanation problem (roughly, a causal-luck or causal-luck hybrid solution). For further discussion of the dialectical significance of the distinct between Pereboom’s “slippery-slope explanation argument” and more standard best-explanation arguments, see Mickelson “Hard Times for Hard Incompatibilism” (manuscript).

³⁵ Kadri Vihvelin has rightly noted that possibilism was a standard background assumption of much of the 20th century free-will debate, and this assumption has serious implications for the dialectic of the free-will debate (Vihvelin 2013, Ch. 2). See also Clarke 2003 and Mickelson *forthcoming*.

³⁶ A detailed discussion of the idea that a libertarianism-friendly account of free will must be assumed in any viable defense of incompatibilism is developed in Mickelson 2019a and “Hard Times for Hard Incompatibilism” (manuscript).

accounts of free will? Does this discussion of the Zygote Argument have connections to the similar on-going disputes about the Consequence Argument, e.g. that it is question-begging (cf. Fischer and Pendergraft 2013) and that it may not (contrary to popular billing) support *any* specific solution to (E2) of the explanation problem (cf. Campbell 2007: 111)? These fruitful new questions are raised by the explanatory gap objection, and their answers are by no means obvious. Since De Marco aims to trivialize the explanatory gap objection when he claims that it easily answered and rests upon distinctions that are not worth “fussing” over, he overlooks the interesting questions and fruitful new lines of inquiry that this objection have opened up.

6 Conclusion

In this essay, I have argued that both of De Marco’s proposed solutions to my explanatory gap objection are nonstarters. De Marco’s “initial solution” describes a successful premise-preserving solution to the objection but it cannot be credited to him: he merely repeats—using his own preferred and arguably problematic jargon³⁷—a type of premise-preserving solution that I had already identified and Mele had already adopted. De Marco’s second solution consists in an invalid variant of the Zygote Argument which, like Mele’s original argument, is subject to the explanatory gap objection. Finally, De Marco failed to see that best-explanation reasoning plays a covert role in making his new Zygote Argument (ZAP-2016) appear valid. This means that De Marco fails to challenge my claim that there is no (non-question-begging way) to close the gap between an impossibility solution to the correlation problem and a specific solution (e.g. a broadly causal luck solution) to (E2) of the explanation problem without using best-explanation reasoning.

If closing the explanatory gap between mere impossibility views and views about what explains/accounts for that impossibility were as trivial and philosophically uninteresting as many free-will specialists appear to think, then why is *fully closing the gap* between an impossibility solution to the correlation problem and a broadly causal luck solution to (E2) of the explanation problem proving to be so difficult? This question remains—along with many other interesting questions raised by the explanatory gap objection—unanswered. In my assessment, such questions help us to see the project of closing the explanatory gap for what it is: a surprisingly difficult and potentially illuminating philosophical challenge.

Acknowledgements I would like to thank the members of the Gothenburg Responsibility Project Reading Group (especially Gunnar Björnsson, Robert Hartman, Sophia Jeppson, Per-Erik Milam, Benjamin Matheson, and Christian Munthe), Keith Lehrer, for helpful comments. For funding, I am grateful to the Swedish Research Council, the Gothenburg Responsibility Project, and the Department of Philosophy, Linguistics, and Theory of Science at the University of Gothenburg.

³⁷ De Marco suggests that the benefit of his relabeling strategy is that it allows philosophers to avoid the “fuss” over the impossibilism/incompatibilism distinction and the terminology that tracks it (De Marco 2016: 1624). In saying this, De Marco is proposing (presumably without recognizing it) that philosophers should not fuss over the distinction between perhaps spurious correlation relations and candidate explanations for those correlations, nor should they fuss over the type of *cum hoc, ergo propter hoc* reasoning that is evident in ZA-2012, nor should they fuss over the difference between an impossibility solution to the correlation problem and the manifold of distinct explanations for that impossibility that are evident in extant solutions to (E2) of the explanation problem.

Funding Open access funding provided by University of Gothenburg.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Anderson, M. B. (2019). Moral luck as moral lack of control. *Southern Journal of Philosophy*, 57, 5–29.
- Bailey, A. (2012). Incompatibilism and the past. *Philosophy and Phenomenological Research*, 85(2), 351–376.
- Campbell, J. (2007). Free will and the necessity of the past. *Analysis*, 67, 105–111.
- Clarke, R. (2003). *Libertarian accounts of free will*. New York: Oxford University Press.
- De Marco, G. (2016). Rescuing the zygote argument. *Philosophical Studies*, 173, 1621–1628.
- Fischer, J. M., & Pendergraft, G. (2013). Does the consequence argument beg the question? *Philosophical Studies*, 166, 575–595. <https://doi.org/10.1007/s11098-012-0053-y>.
- Hartman, R. J. (2017). *In defense of moral luck: Why luck often affects praiseworthiness and blameworthiness*. New York: Routledge University Press.
- Latus, A. (2001). Moral Luck. In *The Internet Encyclopedia of Philosophy*, J. Feiser. (ed.) Available online: <<https://www.iep.utm.edu/moralluc/>> (accessed May 2019). ISSN 2161–0002.
- 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. <https://doi.org/10.1080/00048400903233811>.
- McKenna, M., & Pereboom, D. (2016). *Free will: A contemporary introduction*. New York: Routledge.
- Mele, A. (1995). *Autonomous Agents: From Self-Control to Autonomy*. New York: Oxford University Press.
- Mele, A. (2005). A Critique of Pereboom's 'Four-Case Argument' for Incompatibilism. *Analysis*, 65(1), 75–80.
- Mele, A. (2006). *Free Will and Luck*. New York: Oxford University Press.
- Mele, A. (2008). Manipulation and Moral Responsibility. *The Journal of Ethics*, 12(3), 263–286.
- Mele, A. (2012). Manipulation, Moral Responsibility, and Bullet Biting. Presentation at the *Workshop on The Manipulation Argument*, Central European University (Budapest, Hungary); <https://philosophy.ceu.edu/events/2012-06-07/workshop-manipulation-argument>
- Mele, A. (2013). Manipulation, Moral Responsibility, and Bullet Biting. *The Journal of Ethics*, 17(3), 167–184.
- Mele, A. (2017). *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*. New York: Oxford University Press.
- Mele, A. (2019). *Manipulated agents: A window to moral responsibility*. Oxford: Oxford University Press.
- Mickelson, K. M. (2015a). A critique of Vihvelin's Threefold Classification. *Canadian Journal of Philosophy*, 45(1), 85–99.
- Mickelson, K. M. (2015b). The Zygote Argument is Invalid—Now What? *Philosophical Studies*, 11, 2911–2929.
- Mickelson, K. M. (2017). The Manipulation Argument. In M. Griffith, N. Levy, & K. Timpe (Eds.), *The Routledge Companion to Free Will* (pp. 166–178). New York: Routledge.
- Mickelson, K. M. (2019a). The Problem of Free Will and Determinism: An Abductive Approach. *Social Philosophy and Policy*, 36(1), 154–172.
- Mickelson, K. M. (2019b). Free Will, Self-Creation, and the Paradox of Moral Luck. *Midwest Studies in Philosophy*, 43(1), 224–256.
- Mickelson, K. M. (forthcoming). (In)compatibilism. In *A Companion to Free Will*, Joseph Campbell, Kristin M. Mickelson, and V. Alan White (eds.). New York: Wiley-Blackwell.

- Nagel, T. (1986). *The view from nowhere*. Oxford: Oxford University Press.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
- Sartorio, C. (2015). The Problem of Determinism and Free Will Is Not the Problem of Determinism and Free Will. In A. Mele (Ed.), *Surrounding Free Will: Philosophy, Psychology, Neuroscience* (pp. 255–273). Oxford: Oxford University Press.
- Sartorio, C. (2016). *Causation and free will*. Oxford: Oxford University Press.
- Statman, D. (2019). The definition of luck and the problem of moral luck. In I. M. Church & R. J. Hartman (Eds.), *The Routledge Handbook of the Philosophy and Psychology of Luck* (pp. 195–205). New York: Routledge.
- Stone, J. (1998). Free will as a gift from God: A new compatibilism. *Philosophical Studies* 92(3), 257–281.
- Strawson, G. (1994). The Impossibility of Moral Responsibility. *Philosophical Studies*, 75, 5–24.
- van Inwagen, P. (1983). *An Essay on free will*. Oxford: Clarendon Press.
- Vihvelin, K. (2008). Compatibilism, incompatibilism, and impossibilism. In Theodore Sider, John Hawthorne & Dean W. Zimmerman (Eds.), *Contemporary Debates in Metaphysics* (pp. 303–318). Blackwell.
- Vihvelin, K. (2013). *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. New York: Oxford University Press.
- Vihvelin, K. (Fall 2018). Arguments for Incompatibilism. In *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.); URL = <<https://plato.stanford.edu/archives/fall2015/entries/incompatibilism-arguments/>>.
- Williams, B., and Nagel, T. (1976). Moral Luck. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 50, 15–135 + 137–151.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



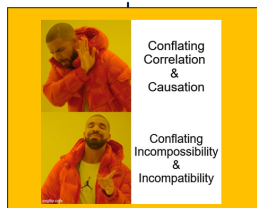
INCOMPOSSIBILITY INCOMPATIBILITY

DET → R1 DET → R2 FW

INCOMPATIBILISM

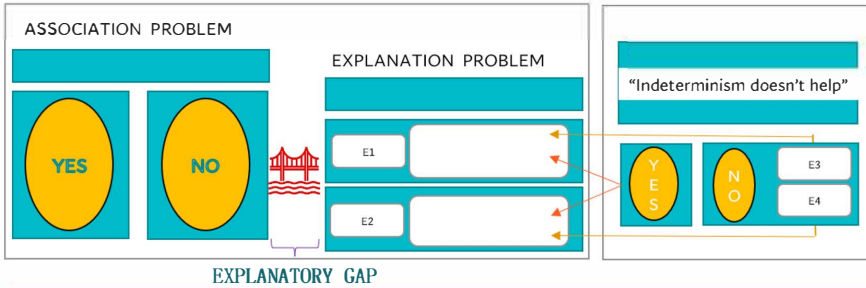
the claim that compatibilism is false. Implementing this strategy would allow the arguments to retain their dialectical role without having to fuss over the difference between incompatibilism and impossibilism.

RELEVANT QUESTIONS



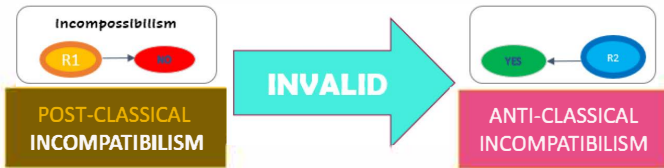
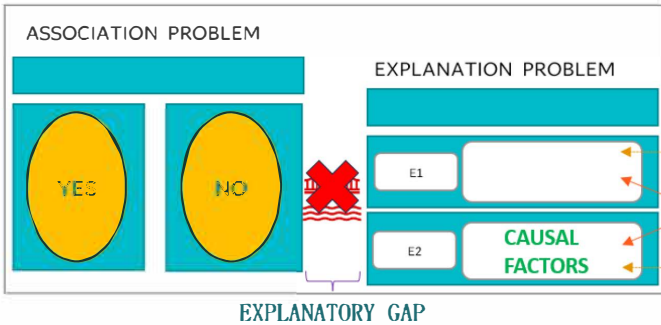
PROBLEM OF DETERMINISM

PROBLEM OF INDETERMINISM

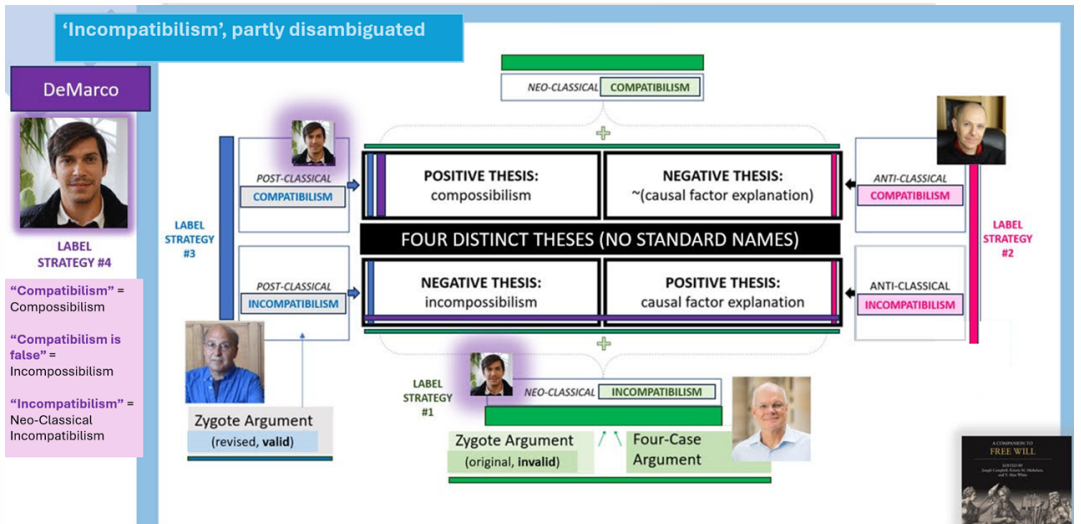


The Paradox of (In)determinism

PROBLEM OF DETERMINISM



'Incompatibilism', partly disambiguated



Mele on 4CA-2001

[=4CA-2022]

“Pereboom has failed to justify his **diagnosis** of the source intuition that Plum is not morally responsible for killing of manipulation: **the deterministic aspect of those cases I have explained**. Consequently, he has failed to provide for his theses that Plum is not morally responsible case 4 and that this is so **because** ‘his action results from a **deterministic causal process** that traces back to factors beyond his control.’

Pereboom's four-case *argument for incompatibilism* fails.”

(Mele 2005: 80, my emphasis).

MELE 2005:

“Pereboom has failed to justify his diagnosis of the source intuition that Plum is not morally responsible for killing of manipulation: the deterministic aspect of those cases I have explained. Consequently, he has failed to provide for his theses that Plum is not morally responsible case 4 and that this is so because ‘his action results from a deterministic causal process that traces back to factors beyond his control.’

Pereboom's four-case *argument for incompatibilism* fails.

So holding fixed the definition of ‘incompatibilism’

Mele

used in his

al dente reply to 4CA-2001,

MY

al dente reply to

OZA-2006 / OZA-2012 shows that:



Mele's Original Zygote Argument for *incompatibilism* fails.



“For **nontraditional** uses of ‘compatibilism’ and ‘incompatibilism,’ see

The Zygote Argument is invalid: Now what?

Kristin Mickelson

-Mele 2017: 6, n. 4

What do crickets



POST-CLASSICAL
INCOMPATIBILISM

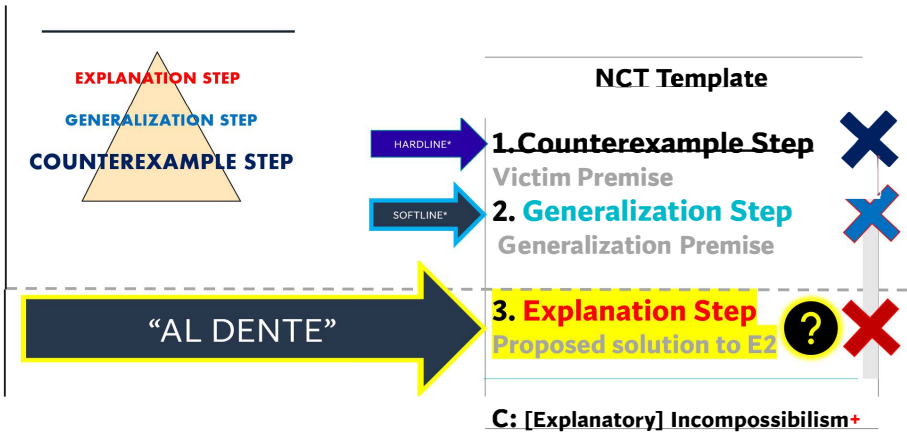
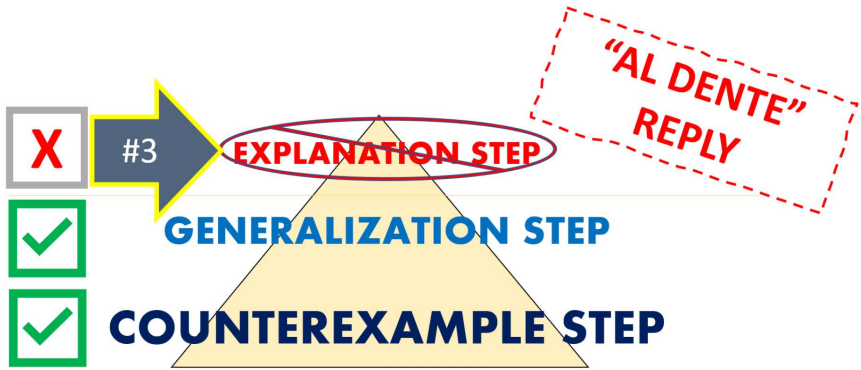
NEO-CLASSICAL
INCOMPATIBILISM

corollary

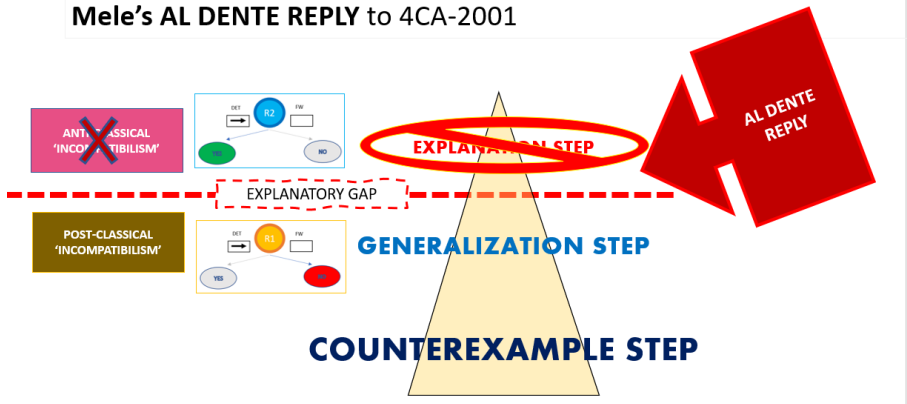
ANTI-CLASSICAL
INCOMPATIBILISM



EXPLANATORY GAP OBJECTION
to a Manipulation Argument:



My **AL DENTE REPLY** to OZA
is **structurally the same as**
Mele's **AL DENTE REPLY** to 4CA-2001



2017

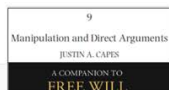
4. For nontraditional uses of "compatibilism" and "incompatibilism," see Mickelson 2015.



JUSTIN A. CAPES 2023: 145

1. *No-responsibility-premise*: the manipulated agent in this case didn't A of his own free will and so isn't morally responsible for A-ing.
2. *No-difference-premise*: when it comes to free will and moral responsibility, there is no relevant difference between the manipulated agent in this case and ordinary, non-manipulated agents in fully deterministic settings.
3. Hence, agents in fully deterministic settings don't act freely and so aren't morally responsible for their behavior.

The argument is valid, but how plausible are its premises?⁴



1. Mickelson did not say this type of argument (for impossibilism) is invalid.

2. 'Incompatibilism' isn't in the premises or conclusion, so its definition is irrelevant to the argument's form



4 Mickelson (2015) contends that arguments like this are invalid. But that's because she defines "incompatibilism" differently than I've defined it here.

Capes also suggests reading De Marco (2016), but De Marco agrees that Mele's argument is INVALID for the reasons that Mickelson gives.

For those manipulation arguments that are invalid for the reasons that Mickelson gives [i.e., those that are subject to an explanatory gap objection], as well as for the Zygote Argument, an easy fix would be to change their conclusion to the claim that *compatibilism is false*. (De Marco 2016:1624, my emphasis)

2 The original Zygote Argument is invalid

2913

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 premise 1 nor 2 identifies deterministic laws as "menacing," i.e. free 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 determinism 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 the zygote was *produced*. However, Mele's description of ZA makes it clear that Ernie should not read premise 1 as promoting a particular account of Ernie's lack of freedom and responsibility. Mele explicitly *denies* that his argument

Philos Stud (2015) 172:2911–2929
DOI 10.1007/s11098-015-0449-6

The Zygote Argument is invalid: Now what?

Kristin Mickelson

What I actually said...

INVALID: association doesn't imply relevance

1. Incompatibilism
2. So, impossibilism is *true in virtue of the antagonistic relationship between determinism and free will*

2914

K. Mickelson

"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 "zeros 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 impossible 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.

Subject: Your Submission

Dear Professor Mickelson,

With regret, I must inform you

PHILSTUDIES DECISION: **REJECT**

2017, 2018, & 2022

A further question is whether the dispute between Mickelson and DeMarco merits a paper in Phil Studies devoted to adjudicating their differences. Mickelson's thesis that Mele's Zygote Argument is invalid already rested on a contentious reading of Mele's Argument to begin with, and even if Mele's argument is strictly speaking invalid, it seems that a very simple remedy would do. (I suppose this is one of the points DeMarco made in his paper.)



All Content

Images

Search journals, books, images, and primary sources



The Zygote Argument is invalid: Now what?

Kristin Mickelson

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 to a

Philos Stud (2015) 172:2911–2929
DOI 10.1007/s11098-015-0449-6

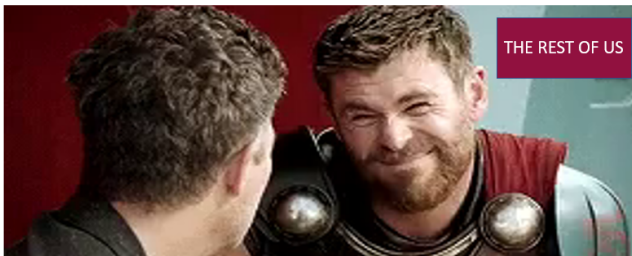


The Zygote Argument is invalid: Now what?

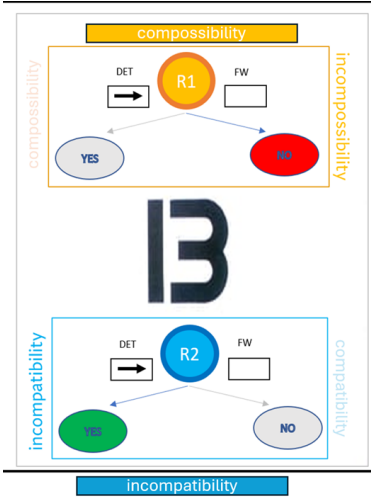
Kristin Mickelson

reading of Mele's Argument to begin with, and strictly speaking invalid, it seems that a very simple remedy would do. (I suppose this is one of the points DeMarco made in his paper.)

The *PRECLUDES* relation is just a symmetrical, negative, & perhaps trivially spurious correlation relation.



IS IT THOUGH ?



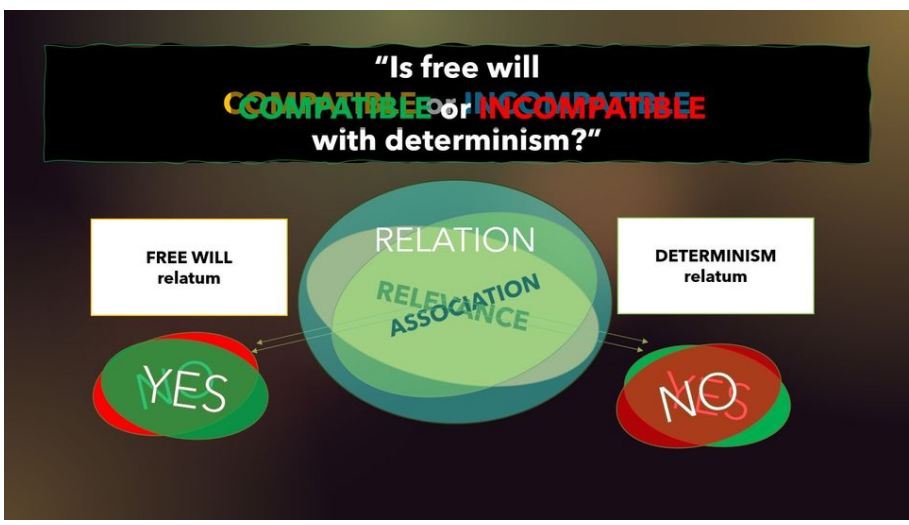
- INCOMPATIBLE
- INCOMPATIBLE
- PRECLUDES
- PRECLUDES
- RULES OUT
- RULES OUT
- THREATENS
- THREATENS
- VITIATES
- VITIATES
- DESTROYS
- DESTROYS
- UNDERMINES
- UNDERMINES
- CONFLICTS WITH
- CONFLICTS WITH
- MAKES IMPOSSIBLE
- MAKES IMPOSSIBLE
- ...
- ...

AMBIGUITY CREEP

Determinism...

- *conflicts with* (Kane 2005: 23; Chisholm 1964/2009: 24)
- *rules out* (McKenna and Pereboom 2016: 169)
- *undermines* (Nahmias et al. 2006: 40)
- *is at odds with* (Pereboom 2005: 240)
- *precludes* (Clarke 2003: 8; Mele 2006: 189)
- *cannot be reconciled with* (Dennett and Caruso 2021: 5)
- *vitiates* (Fischer 1994: 159)
- *would deprive us of* (Vihvelin 2022 & Mele 2002)
- *is menacing to* (Haji and Cuypers 2006)
- *is a problem for* (Campbell 2011: 23)
- *threatens* (Baker 2006: 313)

...Free Will.



Original Zygote Argument (OZA-2006 / OZA-2012)

C. ∴ "INCOMPATIBILISM"

Determinism is

- Incompatible with
- Precludes
- Rules Out
- Undermines
- Destroys

free will (at least for ordinary humans)

B
A
I
L
E
Y

B

B

Revised Zygote Argument (RZA-2012 / RZA-2015)

C. ∴ "INCOMPATIBILISM"

Determinism is

- Incompatible with
- Precludes
- Rules Out
- Undermines
- Destroys

free will (at least for ordinary humans)



M
O
T
T
E

13

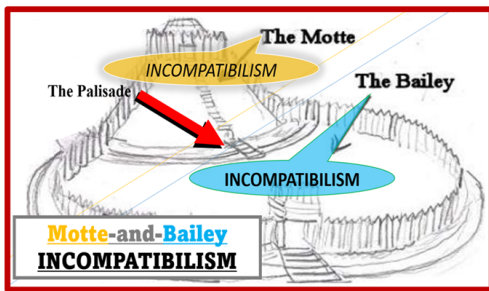
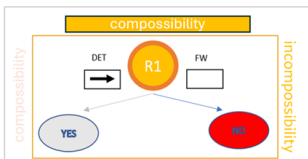
B 13 B 13 B 13 B 13

Explanatory Gap Objection!
Bait-and-Switch!
Equivocation!
Poisoning the Well!
Fallacious Appeal to Authority!
Cum hoc Ergo Propter Hoc Fallacy!

Motte-and-Bailey!

EXPLANATORY GAP

Motte-and-Bailey
INCOMPATIBILISM



Original Zygote Argument (OZA-2006 / OZA-2012)

C. ∴ "INCOMPATIBILISM"

Determinism is

- Incompatible with
- Precludes
- Rules Out
- Undermines
- Destroys

free will (at least for ordinary humans)

B
A
I
L
E
Y

B

Revised Zygote Argument (RZA-2012 / RZA-2015)

C. ∴ "INCOMPATIBILISM"

Determinism is

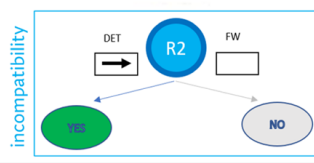
- Incompatible with
- Precludes
- Rules Out
- Undermines
- Destroys

free will (at least for ordinary humans)

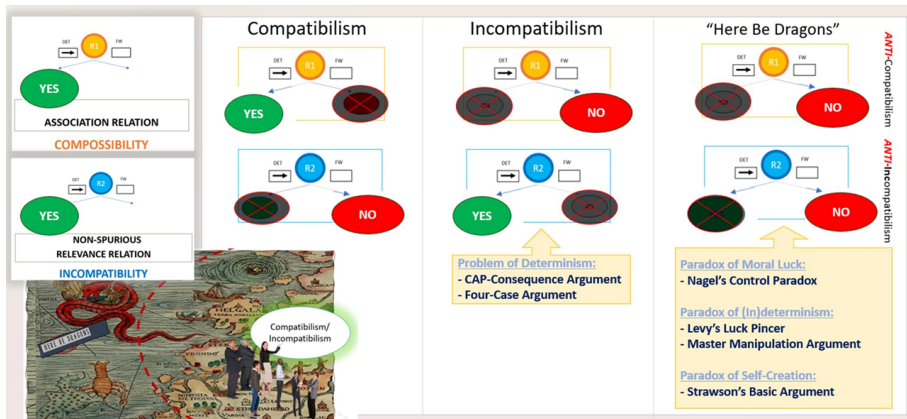
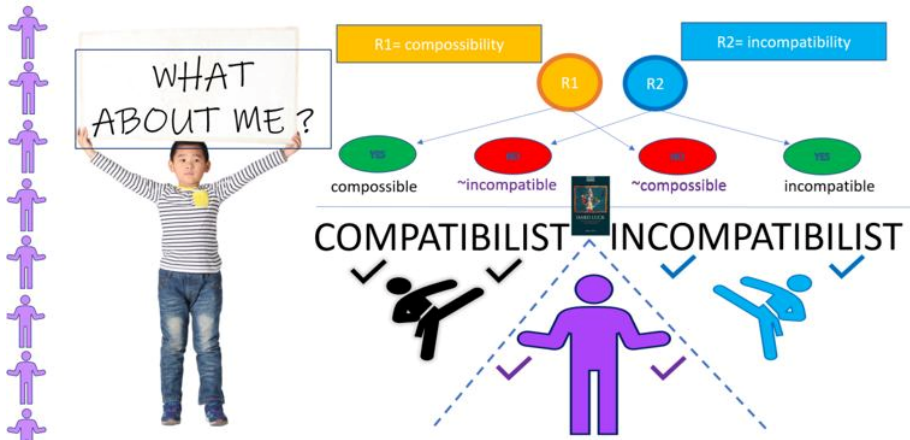
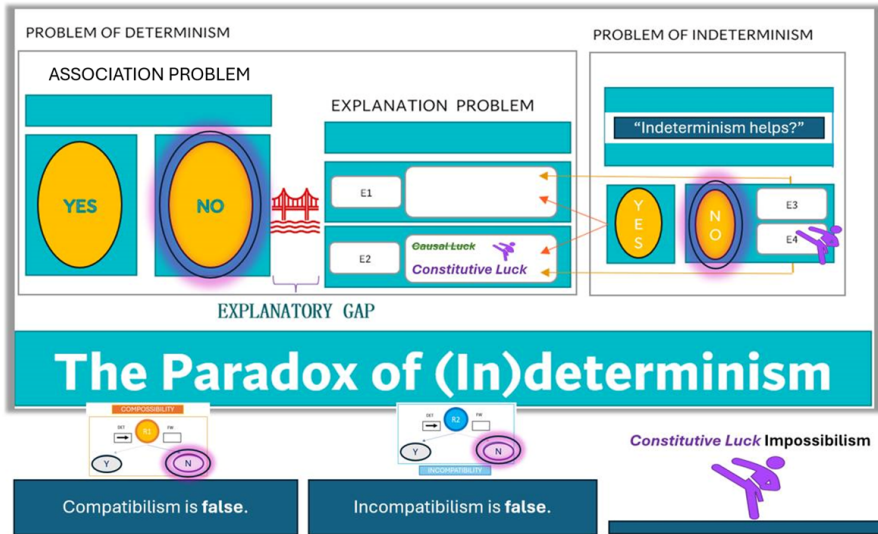


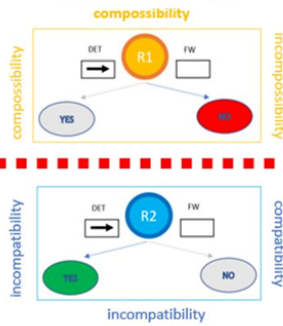
M
O
T
T
E

13



incompatibility





PROBLEM OF DETERMINISM

CORRELATION PROBLEM	EXPLANATION PROBLEM	PROBLEM OF INDETERMINISM
<div style="display: flex; justify-content: space-around;"> <div style="background-color: yellow; border-radius: 50%; width: 20px; height: 20px; display: flex; align-items: center; justify-content: center; font-size: 8px;">YES</div> <div style="background-color: orange; border-radius: 50%; width: 20px; height: 20px; display: flex; align-items: center; justify-content: center; font-size: 8px;">NO</div> </div>	<div style="display: flex; justify-content: space-around;"> <div style="border: 1px solid black; width: 20px; height: 10px; margin-bottom: 2px;"></div> <div style="border: 1px solid black; width: 20px; height: 10px; margin-bottom: 2px;"></div> </div>	<div style="display: flex; justify-content: space-around;"> <div style="background-color: yellow; border-radius: 50%; width: 20px; height: 20px; display: flex; align-items: center; justify-content: center; font-size: 8px;">YES</div> <div style="background-color: orange; border-radius: 50%; width: 20px; height: 20px; display: flex; align-items: center; justify-content: center; font-size: 8px;">NO</div> </div>
EXPLANATORY GAP		

The Paradox of (In)determinism

THE PARADOX PARADIGM

PARADOX PARADIGM

