A COMPANION TO
FREE WILL

EDITED BY
Joseph Campbell, Kristin M. Mickelson,
and V. Alan White

WILEY Blackwell

A Companion to Free Will

Edited by
Joseph Campbell,
Kristin M. Mickelson, and
V. Alan White

WILEY Blackwell

Contents

1

Introduction, *Wiley Companion to Free Will*

KRISTIN M. MICKELSON, JOSEPH CAMPBELL,
AND V. ALAN WHITE

Brackets are used to refer to chapters in this volume (e.g. [10] refers to Chapter 10 of the volume). A glossary of **bolded** terms is provided at the end of this chapter.

We wish this volume to be a sure companion to the study of free will, broadly construed to include action theory, moral and legal responsibility, and cohort studies feathering off into adjacent fields in the liberal arts and sciences. In addition to general coverage of the discipline, this volume attempts a more challenging and complementary accompaniment to many familiar narratives about free will. In order to map out some directions such accompaniment will take, in this introduction we anchor the thirty contributions to this volume in some common history from which they arise, and attempt to indicate where future work in free will and moral responsibility will–and has already begun to–depart from that history.

1 Preliminaries: Free Will and Determinism

The concept of free will is fraught with controversy, as readers of this volume likely know. Philosophers disagree about what free will is, whether we have it, what mitigates or destroys it, and what (if anything) it's good for. Indeed, philosophers even disagree about how to fix the referent of the term 'free will' for purposes of describing and exploring these disagreements (Nichols [28]). What one person considers a reasonably neutral working definition of 'free will' is often considered question-begging or otherwise misguided by another. Such disputes make it difficult to summarize the **problem of free will**, roughly the debate over the nature and existence of free will, in a clear and uncontentious way. In generic terms, however, the two basic solutions to the problem of free will are **free-willism**, the view that we (ordinary humans) have free will, and **free-will denialism**, i.e. the view that we do not have free will (Smilansky [12]).[1] As stated here, neither denialism nor free-willism constitutes a complete solution to the problem of free will; to be complete, a proposed solution must also tell us a convincing story about what free will *is* (what 'free will' *means*) and that, as it turns out, is a very difficult task indeed.

A Companion to Free Will, First Edition. Edited by Joseph Campbell, Kristin M. Mickelson, and V. Alan White.

One historically popular way of approaching the problem of free will is to ask about the *relationship* between free will and determinism: "Does free will stand in relation *R* to determinism: *yes* or *no*?" This is just a template for a question, of course. To transform this template-question into a substantive question with a clear meaning, we need to flesh out the template's free-will relatum, its determinism relatum, and give a precise value to relation *R*. There is, however, no uncontroversial way to do this. In addition to the standard difficulties raised by fixing the referent of 'free will', philosophers hold radically different views about what is – or should be – meant by the term 'determinism' (White [3], Vihvelin[14]; see also Beebee and Mele 2002, Shabo 2010), and they identify relations which are as substantively different as correlation and causation when characterising relation *R* (Mickelson [4]). In practical terms, then, it may be best to think of **the problem of determinism** as a loose collection of disagreements about how to best spell out and answer the template-question, and how (if at all) asking and answering such questions would help us to solve the problem of free will.

The term 'determinism' was ushered into the free-will literature in the 19th century, but the doctrine may be traced back to the Stoic's naturalistic cause-and-effect theory of fate (Bobzien 1998, 2021), which may be contrasted with logical and theological varieties of fate which have also been of traditional interest vis-a-vis free will (Finch [2]). William James, in his influential "Dilemma of Determinism," tells his audience that "no ambiguities hang about this word [determinism] or about its opposite, indeterminism" (James 1884). According to James, determinism "professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be" such that the "future has no ambiguous possibilities hidden in its womb." Indeterminism, says James, is true whenever "the parts have a certain amount of loose play on one another, so that the laying down of one of them does not necessarily determine what the others shall be" (James 1884). Put another way, **traditional determinism** (i.e. determinism as it was traditionally conceived within the free-will debate) is the doctrine that there is a causal or nomological necessity in nature which makes one unique future inevitable given what preceded it. **Traditional indeterminism** is the negation of traditional determinism; it is true if and only if it is false that one unique future is inevitable relative to any arbitrary moment in time (holding fixed the naturalistic factors which account for the evolution of the physical universe and the facts of the past, if any, relative to that time) (e.g. van Inwagen 1990, p. 277). Hereafter, we use 'determinism' as shorthand for traditional determinism and 'indeterminism' as shorthand for traditional indeterminism, unless stated otherwise.

Faced with the idea of determinism, many people, especially those working within the Christian tradition (Adams [7]), have argued that no one could exercise free will in a world at which this necessity-in-nature doctrine of predetermination is true; others–including the ancient Stoics–have disagreed. One popular way of tracking this age-old dispute has been to divide philosophers based on their answers to two questions: (1) "Is determinism true?" and (2) "Do we–ordinary humans–have free will?". Those who answered "yes" to the first question were classified as *determinists*, and they were subdivided based on their preferred response to the second question. Determinists who answered "no" were classified as *hard* determinists, while those who answered "yes" were classified as *soft* determinists (James 1884). The determinists were contrasted primarily with *libertarians*, i.e. philosophers who answered "no" to the first question and "yes" to the second.[2] The term 'hard' in hard determinism indicates that some species of denialism is true. Despite their substantive differences, the soft determinists and libertarians agreed that denialism is false, which is to say that they agreed that free-willism is true.

In an innovative move, a group of philosophers working in the so-called *classical period* of the free-will debate, c. 1965–1985 (van Inwagen 2017), shifted the focal point away from the question of whether determinism is true to more theoretical questions which (according to their diagnosis) lay just under the surface of the pre-classical taxonomy of free-will views. One question was singled out as particularly important:

> Is there a *conflict* or *tension* between the very notions of *free will* and *determinism*, such that *if* determinism were true, it would follow that determinism-related factors, i.e. the causal and/or nomological factors described by determinism, *preclude* free will (as the hard determinists and libertarians propose) or is there no such conflict (as the soft determinists believed)?

To raise the same question in slightly different terms, we could say–following the popular practice of using "luck" as shorthand for *factors beyond one's control* (Hartman [23], Mickelson [4])–that these theorists were focused on the narrow question of whether or not *determinism-related causal luck* poses a distinct threat to free will. The challenge of answering the question of whether determinism (i.e. determinism-related causal factors beyond one's control, determinism-related causal luck) precludes human free will is now widely known as **The Compatibility Problem**.

The Compatibility Problem was initially nested within the dominant research paradigm of the classical period: the **classical analytic paradigm (CAP)** (Mickelson [4]). Among CAP's defining background assumptions, the assumption of **classical possibilism** is especially significant. Classical possibilism may be understood as the conjunction of two claims: (1) the **classical account of free will** is correct, i.e. free will is (some kind of) an ability to do otherwise, and (2) **anthropocentric possibilism**, the view that it is metaphysically possible for an ordinary human to exercise free will, is true (e.g. van Inwagen 1983, Clarke 2003, Vihvelin 2013, Mickelson [4]).

By assuming classical possibilism, CAP theorists (i.e. philosophers working within CAP) restricted the compatibility problem to the **classical compatibility problem**, roughly a debate about which possibilist interpretation of the ability to do otherwise is best. On the one hand, **classical incompatibilists** (e.g. Peter van Inwagen 1983) contend that a person exercises the ability to do otherwise (a.k.a. free will) when performing an action only if there is some kind of *indeterministic leeway* in the evolution of the physical world (see Smilanksy [12], Balaguer [19]).[3] Since all classical theorists accept classical possibilism, classical incompatibilism comes bundled with the endorsement of a broadly libertarian account of free will.[4] As such, it is easy for the classical incompatibilist to explain why there is a deep conceptual *conflict* or an antagonistic *incompatibility* relation between the notions of free will and determinism. Since determinism states that there are naturalistic factors (i.e. determinism-related factors) which eliminate all indeterministic leeway in the evolution of the world, to say that determinism is true is to say that the world includes factors which eliminate the type of indeterministic leeway (whatever type that may be) that an exercise of free will requires.[5] On the other hand, **classical compatibilists** (e.g. Keith Lehrer 1990) argue that indeterministic leeway is not required to exercise the ability to do otherwise. According to the classical compatibilist, the mere fact that determinism-related factors rule out all indeterministic leeway does not mean–*pace* the classical incompatibilists–that determinism-related factors rule out the ability to do otherwise (a.k.a. free will). However, classical compatibilists do not merely reject the classical incompatibilists' "causal factors" *explanation* for the purported fact that normal humans cannot act freely when determinism; they are also committed, given their assumption of classical possibilism, to a classical version of **compossibilism**, the view that it is metaphysically possible

for an ordinary human to act freely in a world at which determinism is true (Mickelson [4]). As such, the classical compatibility problem may be summarized as the challenge of settling which of two views, classical (compossibilist) compatibilism or classical (broadly libertarian) incompatibilism, is correct.[6]

Challenges to CAP have given rise to other perspectives on the compatibility problem and to fundamentally different interpretations of the problem of determinism. Three challenges are worth noting, given their profound impact on the trajectory of the recent history of the free-will debate. Two of these challenges target the CAP assumption that classical possibilism is true, while the third challenges CAP's implicit practice of framing the problem of determinism as a narrow dispute about the relationship between free will and causal luck.

The first major strike against classical possibilism came in the form of Harry Frankfurt's influential criticisms of the classical (a.k.a. leeway) account of free will (Frankfurt 1969) specifically what are now known as "Frankfurt examples" (Haji [6]). By casting doubt on the classical conception of free will, Frankfurt examples motivated interest in non-classical accounts of free will, especially *sourcehood* accounts (Haji [6], Capes [9]). This, which is perhaps the most well-known critique of CAP's background assumptions, was not considered a fatal flaw in the CAP approach to the problem of determinism. Rather, it led philosophers to think that the term 'free will' should not be narrowly defined to mean "an ability to do otherwise" in generic statements of the compatibility problem. In such contexts, 'free will' should instead be defined in a way that opens dialectical space for a lively debate about which account of free will is correct.

This shift in the working definition of 'free will' led to the popularization of the **neo-classical compatibility problem**, which is (at least superficially) just like the classical compatibility problem except that the term 'free will' is used more broadly. The neo-classical use of 'free will' allows that the classical account of free will may be true, but it also allows that some non-classical account (e.g. a sourcehood account) may be correct. The two recognized solutions to the neo-classical compatibility problem are **neo-classical incompatibilism,** the view that is metaphysically impossible because determinism-related factors undermine free will (neo-classically defined) in worlds at which determinism is true, and **neo-classical compatibilism**, the view that determinism-related factors pose no threat to free will and it is metaphysically possible for an ordinary human to exercise free will (neo-classically defined) in a world at which determinism is true.[7] Criticism of the classical definition of 'free will' also contributed to the centralization of *moral responsibility* in neo-classical and non-classical definitions of the term 'free will' (Haji [6], McCormick [24]), a point that we return to below (Section 4).[8]

As sourcehood accounts became mainstream, they helped to normalize the idea that, *pace* CAP theorists, anthropocentric possibilism may be false. While some source theorists, including Frankfurt, became neo-classical *source compatibilists* (arguing that it is possible for an ordinary human to satisfy the necessary source condition on free will even when determinism is true, from which it follows that determinism-related factors do not always undermine free will), other source theorists became neo-classical *source incompatibilists* (arguing that determinism-related factors preclude free will in virtue of keeping people from satisfying the source condition–as opposed to the classical ability-to-do-otherwise condition) on free will. Some of these neo-classical source incompatibilists, e.g. Derk Pereboom 2001, 2014, were also concerned about apparent threats to free will posed by *indeterministic* causal factors (i.e. indeterministic forms of causal luck). Such concerns led to the emergence of Pereboom's **hard source incompatibilism**, a species of **anthropocentric impossibilism** which claims that it is metaphysically impossible for an ordinary human (i.e. someone like us, as we are here on Earth) to exercise free will on the grounds that, whether determinism is true or false, some kind of causal luck (i.e. causal factors beyond one's control) ensures that no normal

human satisfies the necessary source condition on free will.[9] Since hard source incompatibilism clearly speaks against both tenets of classical possibilism–rejecting both the classical account of free will and the assumption of anthropocentric possibilism–it is a decisively non-CAP position.

To be clear, Pereboom's hard source incompatibilism is not an example of full-blooded **impossibilism**, the unqualified view that it is impossible for anyone–even God (Adams [7]; Leftow [11])–to exercise free will. Pereboom is sympathetic to a broadly agent-causal (as opposed to event-causal) libertarian account of free will (e.g. Pereboom 2001, 2014; Vicens [5]) and this keeps him from endorsing unqualified impossibilism. Hard source incompatibilism is an influential view in part because it promises to provide a complete solution to the problem of free will: the "source" part tells us what free will amounts to and the "hard" part signals its endorsement of denialism. The hard source incompatibilist route to solving the problem of free will is attractive, in part, because it allows its proponents to adopt denialism without taking a stand on the truth-value of determinism.

The growing popularity of source accounts of free will has also raised the profile of philosophers who have been arguing for unqualified **impossibilism**. Among impossibilists (e.g. Galen Stawson 1986, Levy 2011, Mickelson 2019b), some argue for the radically anti-CAP view that the specific factors beyond our control which keep us from acting freely are not located in our environment (e.g. states in the remote past or the laws of nature) but are instead located *entirely* in facts about us. Drawing again on the language of "luck," these source impossibilists contend that causal luck is irrelevant to free will. They claim, instead, that *constitutive luck*–roughly luck in the way that one is constituted, especially in regards to how one is mentally (at least in certain key respects), at the time of action–keeps people from acting freely, no matter what one's environment is like. This **constitutive-luck source impossibilism**, like its rival hard (source) incompatibilism, provides a route to denialism which does not require us to resolve tricky empirical questions about whether determinism is true or false.

Since constitutive-luck source impossibilism is in direct conflict with all three of the CAP tenets discussed above, it is a paradigmatically non-CAP position. It should not be surprising, then, that this view defies classification in CAP-based terms (e.g. Vihvelin 2008, McKenna and Pereboom 2016, p. 151, Mickelson [4]). Since these impossibilists reject the **compossibilist** component of classical/neo-classical compatibilism, they are not "compatibilists" in any traditional sense; but these impossibilists are not "incompatibilists" in the traditional sense either, for they also reject the explanatory tenet of classical/neo-classical incompatibilism which identifies determinism-related factors (i.e. causal luck) as relevant to free will. Just as there is a clear sense in which constitutive-luck source impossibilism is both an *anti-compatibilist* and *anti-incompatibilist* position, there is also a sense in which it is both a compatibilist and incompatibilist position: it is *incompatibilist* insofar as it entails the modest incompossibilist tenet of traditional forms of incompatibilism, but *compatibilist* insofar as it denies that determinism-related factors pose a threat to free will (for further discussion, see Mickelson [4], 2015a, 2019b). Philosophers have yet to reach a consensus on whether–and, if so, how–to update CAP-based jargon so that it tracks non-CAP views.[10]

The chapters in this volume reflect a variety of classical, neo-classical, and non-classical perspectives on the problem of free will and the problem of determinism. While CAP remains a powerful and popular research framework, alternative approaches promise to raise new questions and inspire fruitful lines of inquiry. A solution to the problem of free will may still be far off, but these new developments should help free-will theorists push back against the common complaint that the free-will debate is still mired in a dialectical stalemate between "compatibilists" and "incompatibilists."

2 Compatibility Concerns: The Arguments

From the perspective of CAP theorists, the problem of free will is just the problem of determinism, and the problem of determinism boils down to the question of whether the thesis of determinism is logically incompatible with the **classical free-will thesis**, i.e. the thesis that some ordinary human exercises free will (assuming the classical definition of 'free will') (Mickelson [4]). The worry, in general terms, is that a certain kind of necessity (determinism) is at odds with a kind of contingency (free will). Looked at in this way, CAP compatibility concerns are part of a family of traditional worries raised by **predeterminisms**, including not only well-known problems about determinism (Campbell and Lota [8]), God's omniscience (foreknowledge) and the logical principle of bivalence (Finch [2]), but also eternalism (Buckareff 2019), providential determinism, and socio-economic determinism. Many of these predeterminisms involve commitments to scientific, religious, even political world views. For instance, a Catholic might be committed to providential determinism, or a Marxist to socio-economic determinism. This partly explains why some compatibility problems are worrisome to some people, but not to others. If the predeterminism is disconnected to one's world view, it is easy to give it up. Once we consider compatibility problems broadly–as involving any number of predeterminisms in conflict with free will–it is likely each of us has a worrisome compatibility problem waiting to be revealed.

The problem of determinism remains a popular entry point to the problem of free will, but it is not the only framework which draws upon notions of luck (i.e. factors beyond our control) to raise pressing questions about the nature and existence of free will. Even if one were to show that the future is not perfectly predetermined–by God, the laws of nature, the axioms of logic, or anything else–one would not have thereby made the case for free will. Even if a world without a pre-established future must include some type of indeterminacy, it is by no means obvious which type of indeterminacy is required. This raises a new concern: perhaps the best arguments in the literature, when taken together, will support the conclusion that free will is impossible whether or not there is indeterminacy in the world and, hence, that denialism is true.

From the ancient Epicurean idea that free will might be found in the random "swerve" of Democritean atoms (Pereboom 2009, pp. 17–18) to the modern idea that free will is grounded in the (purportedly) probabilistic behavior of quantum particles (Kane 2003; Balaguer [19]), many people have argued for a tight connection between free will and causal indeterminacy. As we have seen, CAP theorists are committed to solving the problem of free will through a very particular characterization of the problem of determinism and, given their commitment to classical possibilism, classical incompatibilists are committed to a broadly libertarian interpretation of free will. However, even CAP theorists who are committed to a libertarian analysis of the ability to do otherwise respected the worry that causal indeterminacy might "hurt" one's efforts to exercise free will. For example, van Inwagen's "freakish demon" manipulation argument (van Inwagen 1983, pp. 130–134) was the first of the so-called *manipulation arguments* (Capes [9], Mickelson 2017) to raise serious concerns about the incompatibility of free will and *indeterminism*. The more renowned *Mind* argument raised the same concerns (Campbell and Lota [8]). (It is called the "*Mind* argument" because influential versions of it were published in the journal of that name.) The *Mind* argument "occurs in three forms" or "three closely related strands of argument that are often twisted together" (van Inwagen 1983, p. 126). All the strands begin with "a certain set of reflections on what the nature of free action must be if the incompatibilist is right," e.g., supposing the world is causally undetermined but productive of free action.[11]

Van Inwagen notes there are structural similarities underlying the *Mind* Argument and the Consequence Argument, the most influential argument for classical incompatibilism, suggesting that if one is sound, then so must be the other (van Inwagen 1983, pp. 147–150; Campbell and Lota [8]). In broad strokes, the Consequence Argument is a seemingly simple conditional argument: If determinism is true, then everything we will ever do is a consequence of the laws of nature and states of the world in the remote past (prior to the existence of the first human); since we have no control over the past (Wasserman [10]) or the laws (Vihvelin [14]), we have no control–in just the sense picked out by 'free will'–over anything we do. Yet, it seems not to matter much if we replace the laws of nature with probabilistic laws. Either way, we have "precious little free will" (van Inwagen 1989, p. 405).

As a CAP theorist, van Inwagen originally presented the Consequence Argument against the backdrop of CAP's background assumptions, which places constraints on how we interpret this argument's premises and conclusion. For example, the original CAP-based version of the Consequence Argument (hereafter, the *Classical Consequence Argument*) was specifically an argument for classical incompatibilism. That is, the Classical Consequence Argument concludes that it is impossible to exercise the ability to do otherwise picked out by the term 'free will' when determinism is true, from which it follows (given the CAP assumption of classical possibilism) that a libertarian interpretation of the ability to do otherwise must be correct. According to this libertarian account, indeterministic leeway is a prerequisite for exercising the ability to do otherwise, a.k.a. free will.

Here, then, we strike a tension at the core of the CAP program. The Classical Consequence Argument concludes that classical incompatibilism is true, and above we noted that some philosophers believe that there are structural similarities between the Classical Consequence Argument and the *Mind* argument which ensure that if one of these arguments is sound, then so is the other. However, if both of these arguments are sound, it means that people cannot act freely whether determinism is true or false–in which case classical possibilism, a defining background assumption of CAP, is false.[12] The tension may indicate a problem with the assumption of classical possibilism, i.e. perhaps anthropocentric possiblism and/or the classical account of free will is incorrect (Campbell and Lota [8]). Not wanting to give up on such foundational CAP commitments, it is perhaps unsurprising that van Inwagen–an eminent CAP theorist–has responded to the apparent paradox within CAP by adopting *mysterianism*, the view that free will exists but it is a mystery (van Inwagen 1983, 1998, 2000).[13] For those less committed to the CAP program, the best response to this tension may be less clear–though, minimally, it encourages us to explore other (neo-classical and non-classical) options.

The manipulation argument has become one of the most popular tools for exploring non-CAP approaches to the problem of free will (Capes [9]). Multiple-case manipulation arguments were already in play during the classical period, e.g. van Inwagen's "freakish demon" argument targeted broadly libertarian accounts of free will (van Inwagen 1983, pp. 130–134) and Richard Taylor's earlier "puppet" argument targeted compossibilist accounts of free will (Taylor 1963, p. 45). However, manipulation arguments are now used to support a wide variety of conclusions. For example, Derk Pereboom's influential Four-Case Argument aims to establish neo-classical incompatibilism and to offer some support for the more specific source incompatibilist position that determinism-related causal factors preclude human free will by keeping people from satisfying the source condition on free will (Pereboom 2001, 2014). Other manipulation arguments are more thoroughly untethered from the CAP framework. For example, Alfred Mele's revised Zygote Argument (Mele 2013, 2017, 2019) is distinctive insofar as it concludes to mere **incompossibilism**, the relatively modest

non-explanatory claim that it is impossible for an ordinary human to act freely when determinism is true. Unlike Pereboom's Four-Case Argument, Mele's argument is completely silent about *why* incompossibilism is true (Mickelson 2015b, 2017, 2021). Kristin Mickelson's Master Manipulation Argument marks an even more radical departure from the classical program (Mickelson 2019a, 2019b). Like other influential non-classical arguments, such as Galen Strawson's Basic Argument and Neil Levy's related "Luck Pincer" (Levy 2011, Hartman [23]), the Master Manipulation Argument concludes to constitutive-luck source impossibilism–which, if true, would mean that the rival explanatory conclusion of the Four-Case Argument is *false* (Mickelson 2015b, 2017, 2019a, 2019b).[14]

The expansion of arguments and worries about the relationship between free will and factors beyond our control has generated new thoughts about the problem of free will, and lends force to a relatively new type of worry. With the array of views about free will now available, we can ask "Which reflects the layman's notion of free will–and how should we respond if it turns out that there is a *conflict* between the view philosophers think is the best and the one endorsed by the folk?". These and related worries have motivated *revisionism* about free will, the view that the correct solution to the problem of free will clashes with the folk notion of free will and/or common freedom-related practices, e.g. moral praise/blame and punishment (Vargas [13]). While revisionism raises many interesting and pressing questions about free will and metaphilosophical issues facing those who study it, the justification for revisionism will depend largely upon what our best science tells us about its empirical components (e.g. what the folk think about free will).

3 Science and Free Will

From antiquity, many philosophers have viewed the fixity of the world–whether due to gods, causal-like conditions, the principles of logic, or the like–as antithetical to the belief that humans have any control over their lives. When Newtonian physics arose, the specific challenge presented by causal determinism became especially pressing, for it quickly appeared to be foundational for a scientific view of the world. Subsequent centuries strongly reinforced the explanatory and predictive force of Newton's mechanics with expansion of its influence into other developing sciences such as chemistry and biology, and even began to influence the development of modern psychology through Freud and later more explicitly so with Skinner and Watson's behaviorism. On the practical side, the use of Newtonian principles became crucial to emergent technologies exhibited in the industrial revolution, providing forceful everyday evidence of their increasingly plain truth.

However, this high tide of determinism ebbed somewhat in the early 20th century with the rise of an alternative account of fundamental reality: quantum theory and the associated idea of probabilistic causation or even outright indeterminism at work in the deepest levels of reality (at least according to some interpretations). These latter interpretations brought new hope to aspiring libertarians but also raised new worries for those who believed that quantum-like indeterminacy in human nature could do nothing to aid a like account of free will, unequipped with any feature that could easily accommodate room for human control over it. In the light of these more recent scientific trends, it is not surprising then that the determinism/indeterminism debate arose and continued to strongly influence free-will theorists as informed by the constantly evolving results of scientific inquiry and emerging theories. As we related earlier, this conflict between determinism and indeterminism was philosophically sharpened and focused in the 20th century, giving rise to CAP's emphasis on this distinction, and continues in various themes today (see Vihvelin [14] for a detailed contemporary examination of

the concept of causality at work in determinism, for example). It remains to be seen whether CAP or non-CAP perspectives along with further scientific investigation will move us closer to a satisfactory solution to the problem of free will. It is undeniable though that science has been a formative factor in the free will debate in the last century, and more recently has assumed a prominent role in the very methodology of how to conduct that debate.

In the 1980s empirical experiments such as Libet's famously began to lay ground for the still evident stand among many neuroscientists against the existence of free will (assuming that a broadly libertarian account of free will is correct) citing traceable data that unconscious predispositions for choice can be recorded even before any such choices enter the conscious realm (Libet et al. 1983; see also Waller [17]; Levy and Wright [15] examine one facet of this in terms of implicit attitudes). Since these results may be thought to favor determinism of the mind (in some sense) prior to instances of choice, then indeterminism of choice, either conscious or unconscious, would seemingly be ruled out (but see Woodward [16]). Libertarianism thus appears to be completely knocked out of the realm of plausibility (though many following Libet simply then do not make argumentative room for the possibility of a compatibilist/compossibilist view of free will, thus revealing their bias that some broadly libertarian account of free will must be correct; see Cova [18] for discussion). However, many have pointed out that this is a rush to judgment given the uncertainty of what the data really reveals as against several plausible alternatives of how metaphysically choice may work moving from unconscious sources into the arena of how conscious choices are made (Robichaud [21] examines one important aspect of this).

The rise of neuroscience in the latter part of the 20th century also now plays an important–and some say indispensable–role in understanding how free will and action theory issues sort out against the background of studies about the brain and mental behavior ((Waller [17]); on one extreme end Penrose 1989 argues that a quantum theory of brain activity may solve the free will problem and in a way favorable to libertarianism, but see Boolos et al. 1990 for criticism). Indeed, some advocates of libertarianism have insisted that a careful examination of the science of the brain supports that view and speaks against a Libet-style conclusion of his own studies (Balaguer [19]; see also Kane 2003). As we better understand the details of our mental lives in scientific terms, we may discover at least important clues about how to better interpret any role that freedom and responsibility might play out with respect to those findings.

Aside from the determinism/indeterminism debate, the most direct empirical trend of the 21st century involving the free will problem has been in the rise of experimental philosophy – usually termed "X-Phi" (see Nahmias et al. 2005 and Nichols and Knobe 2007 for example; for metacriticism see Cova et al. 2018). The motivation for X-Phi is rooted in familiar claims in previous free will literature (especially the CAP-based distinction between compatibilism and incompatibilism) about the beliefs and attitudes of "the folk."[15] X-Phi developed in part to inform such claims with real data – statistical surveys that posed specific sets of questions to groups of individuals in order to ascertain real-world beliefs and attitudes about matters of freedom and free will. The idea was that if one could obtain real world data about the intuitions of large groups of people about specific free will-related scenarios then one could tabulate in a comprehensive way overall views that then could factor into free will debates, thus eliminating pure speculation about how "the folk" were disposed to talk in favor of a scientific basis for such claims. In addition, these empirical methods have been extended to include methods testing for psychological factors such as implicit bias, which subconsciously could influence conscious decisions of a free will nature (Levy and Wright [15]). While the significance of this overall avenue of inquiry is still controversial, there is little doubt that these empirical methods will have continuing influence in the ways that forthcoming debates on free will are framed.

4 Moral Responsibility

While the relationship between free will and moral responsibility has always been part of the free-will debate, the latter has become even more prominent as CAP's influence has waned. When the classical (ability to do otherwise) characterization of free will was challenged, philosophers generally agreed that a more inclusive definition of 'free will' was needed for purposes of framing the problem of free will and the problem of determinism. The new definition needed to avoid any details that might be considered question-begging (e.g. by presuming that a classical rather than source account of free will is correct, or vice versa), yet it also needed to be adequately precise to pick out a distinct phenomenon as the topic of debate (lest the free-will discourse devolve into an empty verbal dispute). In response, many philosophers have adopted the practice of using 'free will' to refer to the necessary *control condition*–as opposed to the necessary *epistemic condition* (Robichaud [21])–on moral responsibility, where the latter is understood in the backward-looking, non-consequentialist, type of responsibility associated with *basic desert* (McCormick [24]).

Whether or not one approves of the move towards identifying free will with the type of control required for basic-desert moral responsibility, the moral-responsibility turn in the free-will literature has had its benefits. While moral responsibility is interesting in its own right, the neo-classical practice of fixing the referent of 'free will' in terms of moral responsibility has helped us to approach the problem of free will in new ways and encouraged us to reconsider what the free-will debate is and/or should be about (White [31]). For example, it is commonly agreed that free will is a type of *control* that one exercises in the performance of an action, and that anyone with free will would have, at minimum, the type of control required to make a person praiseworthy and/or blameworthy for their morally-valenced actions. As such, free will seems required to make a person an apt target of the *moral emotions* (Ekstrom [22]) and familiar practices of praise and blame. As such, settling what type of control (if any) is really needed for these things may help us to get a better grasp on what a viable solution to the problem of free will must look like. For example, many free-will theorists have been skeptical of the proposal that 'free will' picks out (or should pick out) the type of control required for ultimate "heaven-and-hell" moral responsibility (Strawson 1994, Adams [7]), a type of responsibility implicit in the belief that God will ensure that humans receive their just deserts, e.g. being tormented in hell in the afterlife. Not only does such *ultimate control* seem to be metaphysically impossible or even incoherent (e.g. Strawson 1994, p. 8; van Inwagen 1998, Mickelson 2019b), but some hold that a comparatively modest type of control–perhaps even more modest than basic-desert responsibility (if there is a difference between the two)–would be enough to support our current moral practices of praise/blame (McCormick [24]), forgiveness (Ekstrom [22]), and reward/punishment. If this is right, then perhaps the idea that free will is intimately connected with some type of ultimate or basic-desert control is mistaken. However, if such practices and policies are justified only if we are at least basic-desert responsible for our actions, then free-will denialism would seem to imply that the time has come to revise these and closely related practices and policies (or at least the justification for them), such as legal policies which recommend harsh punishments for criminal behavior (Ekstrom [22], K. Levy [25], Pereboom [29]).

Among the more recently developed moral-responsibility approaches to the problem of free will are those which focus on the moral agency of the mentally disabled and young children (Griffith [20]) and those which draw upon the well-established literature on the *paradox of moral luck* (Hartman [23]). Since the paradox of moral luck emerged during the classical period of the free-will debate (Williams and Nagel 1976; see also Nagel 1986), it is to be expected that

moral-luck theorists have typically assumed the CAP view of the problem of determinism, i.e. that it is a narrow debate about "antecedent causal luck". However, the recent convergence of the basic vocabulary between the two literatures–especially the language of *control* and *luck*–has highlighted hitherto overlooked similarities between the two problems. As the cross-pollination of these established literatures increases, we can expect more critical pressure to fall upon the CAP assumption that the problem of determinism is fundamentally a problem of causal luck–as opposed to, say, a problem of constitutive luck (Mickelson 2019b). Future work which explores these non-classical avenues of thought may prove equally useful to philosophers working on the problem of free will and to those philosophers who are interested in free will only insofar as it is related to moral and legal responsibility.

5 The Future

The wisdom of speculation about the future of anything has considerable history against it. From the supreme confidence in Newtonian physics prior to Einstein and Planck to the declaration that World War I was so horrific as assuredly to constitute "the war to end all wars", there are countless examples of the retrospective frivolity of predicting the future that seem to undercut the wisdom of even the attempt to do so. However, just as the role of hypothesis is central to the work of much science, and has proved its merits time and again even though failures vastly outnumber successes, we believe that some prognostication about the future of free will and action theory might yield some parallel advantage. This is how we approach such an effort in this volume–not only directly trying to predict how things might go in these and related areas (Mele [30], Tierney [27]), but also emphasizing present areas of investigation that might prove much more fruitful in the future.

The major future trend we see in several contributions in this volume is an emphasis on the role of empirical methods in contributing to or even guiding the dialogues on free will and action theory, as we noted above concerning the rise of X-Phi and the increasing influence of neuroscience. Another facet of this empirical dimension to the debates is that there appears to be an increased emphasis on the phenomenology of choice (Deery and Nahmias [26], Robichaud [21], Woodward [16]; also see Mele [30]). How such an introspective factor argumentatively plays off against more traditional empirical treatments of psychological data requires much more investigation.

A separate trend rooted in empiricism is that of a pragmatic approach to the free will problem, which although implicitly present in the literature since at least P.F. Strawson's influential "Freedom and Resentment" (Strawson 1962), has not been overtly promoted as a dominant theme. Revisionism (Vargas [13]) incorporates some trace of this in its relativizing the adequacy of a concept of free will to its overall workability at any given time (see Ekstrom [22] on this as well). Illusionism (Smilansky [12]) is partly pragmatic by conceding the falsity of libertarianism yet arguing that we need such a concept in moral and legal terms in order to best work as societies (see Levy [25] on how free will is incorporated in matters of legality). In this volume it is argued as well that pragmatism yields the best approach to defining key concepts such as determinism (White [3]), and perhaps is the best overall approach to the entire free will problem (White [31]).

Beyond considerations of the empirical, it appears clear that forms of free will denialism, skepticism, and even to an extent illusionism will expand in influence (Hartman [23], Pereboom [29], Smilansky [12]; also see Vilhauer 2012 and Mele [30] who offer varieties of epistemic skepticism about free will). Though some routes to skepticism, denialism (especially

Pereboom's), and illusionism are extensions of the CAP program, others are firmly outside the CAP tradition (e.g. Galen Stawson 1986, Levy 2011, Mickelson 2019b; see also Mickelson [4]). These non-classical approaches have put considerable pressure on the CAP assumption of classical possibilism, and we believe especially that these forms of skepticism and denialism will increase in influence. Non-classical explorations may well have an influence on future developments in X-Phi as well, leading to better inquiries informed by considerations of the roles of luck in our choices and moral lives and perhaps even leading us to see that intuitions favoring impossibilism are more widespread than currently assumed. Such inquiries could lead to an increased pragmatic concern with reforming our more formal and legal blaming practices in society.

Of course, familiar philosophical approaches in the tradition of CAP or in direct/indirect criticism of it will certainly also maintain significant influence in much or most of the literature (in this volume: Campbell and Lota [8], Mickelson [4], Balaguer [19], Adams [7], Nichols [28], Finch [2], Leftow [11], Mele [30], Vicens [5], Pereboom [29], White [31] for example) and it is clear that this is an important part of moving the field forward by the continued reliance on the time-honored methods of logically-constrained speculation. After all, even in science the source of hypothesis is always the rigorous application of the inventive prowess of the human mind to intriguing problems.

Glossary

Anthropocentric Impossibilism: The view that it is metaphysically impossible for an ordinary human to exercise free will. Anthropocentric impossibilism entails **free-will denialism**, but it does not entail **impossibilism**. **Hard incompatibilism** is a species of anthropocentric impossibilism, though it is *not* a species of impossibilism.

Anthropocentric Possibilism: The view that it is metaphysically possible for an ordinary human to exercise free will. Anthropocentric possibilism entails that **anthropocentric impossibilism** is false and that **possibilism** is true; it is silent vis-a-vis the truth-value of the **free-will thesis** and **free-willism**.

Classical (a.k.a Leeway) Account of Free Will (or the classical definition of 'free will'): Free will is an ability to do otherwise; typically contrasted with *source* accounts of free will. Within CAP, the classical account of free will was assumed to be true, leaving open the debate between **classical compatibilists** and **classical incompatibilists** over which interpretation of the ability to do otherwise is correct.

Classical Analytic Paradigm (CAP): The dominant research paradigm during the classical period (c. 1965–1985) of the free-will debate. The terms 'compatibilism' and 'incompatibilism' were coined for use within CAP, and the background assumptions of CAP play an essential role in justifying the familiar CAP narrative that these two terms named the only two viable candidate solutions to the **classical compatibility problem**.

Classical Compatibilism: The CAP–based view that the **classical account of free will** is correct and that it is metaphysically possible for an ordinary human to exercise free will (where 'free will' refers to an ability to do otherwise) when **traditional determinism** is true, i.e. necessarily, determinism is logically compatible with the **classical free-will thesis** (Mickelson [4]). The "classical" qualifier signals that the classical definition of 'free will' is used in stating the view. The term 'compatibilism' was coined (in the 1960s) as a name for this view.

Classical Compatibility Problem: According to CAP theorists (i.e. philosophers who endorse and work within CAP), the **problem of determinism** boils down to the challenge of settling whether **classical compatibilism** or **classical incompatibilism** is true. Given that **classical possibilism** is a background assumption of CAP, all classical compatibilists and classical incompatibilists were **anthropocentric possibilists**.

Classical Free-Will Thesis: The thesis that free will is (or requires) an ability to do otherwise (i.e. the **classical account of free will** is correct) and some ordinary human exercises free will; put another way, the thesis that an ordinary human exercises free will, where 'free will' refers to an ability to do otherwise.

Classical Incompatibilism: The CAP-based view that **the classical account of free will** is correct and it is metaphysically impossible for an ordinary human to exercise free will when **traditional determinism** is true because determinism-related factors preclude the type of indeterministic leeway that an exercise of free will requires, i.e. necessarily, traditional determinism is logically incompatible with the **classical free-willism** (Mickelson [4]). The "classical" qualifier in the name signals that the classical definition of 'free will' is used in stating the view. The term 'incompatibilism' was coined (in the 1960s) as a name for this view. (If classical incompatibilism is true, then so is **neo-classical incompatibilism**, but not vice versa.) The Classical Consequence Argument (i.e. the Consequence Argument, as originally presented against the background of CAP) concludes to classical incompatibilism.

Classical Possibilism: the conjunction of two views: (1) free will is (some kind of) an ability to do otherwise, i.e. the so-called **classical account of free will** is correct, and (2) **anthropocentric possibilism**.

Compossibilism: The view that it is metaphysically possible for an ordinary human to exercise free will in a world at which traditional determinism is true; the conjunction of determinism and the **free-will thesis** is metaphysically possibly true.

Constitutive-luck Source Impossibilism: The view that it is metaphysically impossible for anyone (i.e. any metaphysically possible being) to exercise free will because constitutive luck–as opposed to, say, causal luck–prevents people from satisfying the necessary source condition on free will. Galen Strawson's Basic Argument (Strawson 1994, 2011) and Kristin Mickelson's Master Manipulation Argument (Mickelson 2021) each conclude to this view.

Free-will Denialism: One of two basic solutions to the problem of free will (the other is **free-willism**). Denialism is the view that no (ordinary human) has free will, i.e. the view that **the free-will thesis** is false. Hard determinism is a common route to denialism; all arguments for **anthropocentric impossibilism** and **impossibilism** are (*a fortiori*) arguments for denialism.

Free-will Thesis: The thesis that an ordinary human exercises free will (where 'free will' is neutral between classical and non-classical accounts of free will, e.g. by fixing the referent of 'free will' to the control condition on basic-desert moral responsibility). Compare to the **classical free-will thesis.**

Free-willism: One of two basic solutions to the problem of free will (the other is **free-will denialism**). The view that some ordinary human has free will, i.e. the view that the **free-will thesis** is true. Libertarianism and soft determinism are common species of free-willism.

Hard incompatibilism: A species of **anthropocentric impossibilism** which claims that it is impossible for an ordinary human to exercise free will on the grounds that, whether determinism is true or false, some kind of causal luck (i.e. causal factors beyond one's control) would keep a normal human from satisfying some necessary condition on free will. The "hard" in the name signals that the view entails **free-will denialism**. Notably, hard incompatibilism is a species of **anthropocentric impossibilism** but not (unqualified) **impossibilism**.

Hard Source Incompatibilism: The view that **hard incompatibilism** is true, and the necessary condition which an ordinary human cannot satisfy when determinism is true is a *source* condition and not a classical ability-to-do-otherwise (a.k.a. leeway) condition.

Impossibilism: The unqualified view that it is metaphysically impossible for anyone (i.e. any metaphysically possible being, even God) to exercise free will. Impossibilism entails **denialism**. Galen Strawson's "Basic Argument" (Strawson 1994, 2011) and Kristin Mickelson's "Master Manipulation Argument" (Mickelson 2021) conclude to impossibilism. (Notably, **hard incompatibilism** is not an impossibilist view.)

Incompatibilism: The term 'incompatibilism' has become an umbrella term and currently has no standard meaning; the same is true of many phrases commonly associated with this term, e.g. "free will is incompatible with determinism". The term is currently used to refer to **incompossibilism, classical incompatibilism, neo-classical incompatibilism**, **anthropocentric impossibilism**, **impossibilism**, and many other views (e.g. see endnote 10). (The same is true, *mutatis mutandis*, of the term 'compatibilism' and the ambiguous phrases commonly used to define it).

Incompossibilism: The view that it is metaphysically impossible for an ordinary human to exercise free will in a world (or universe) at which **traditional determinism** is true; alternatively, the view

that the material conditional "If traditional determinism is true, then the **free-will thesis** is false" is necessarily true (i.e. true in all possible worlds) (for more detail, see Mickelson [4]). Some philosophers now use 'incompatibilism' to refer narrowly to incompossibilism (e.g. Mele [30], Capes [9]; see endnote 10 for discussion). Within CAP, any argument for incompossibilism was also (given CAP background assumptions) an argument for classical incompatibilism; outside of CAP, the inference from incompossibilism to classical incompatibilism or neo-classical incompatibilism is a fallacious *cum hoc, ergo propter hoc* ("with this, therefore on account of/because of this") inference (see Mickelson [4] and 2021). Alfred Mele's revised Zygote Argument (Mele 2013, 2017, 2019; Mickelson 2015b) is an example of an argument for mere incompossibilism.

Neo-Classical Compatibilism: The view that it is metaphysically possible for an ordinary human to exercise free will when **traditional determinism** is true, and traditional determinism does not stand in any antagonistic relevance relation to free will (where 'free will' is neo-classically characterized in a way that is neutral between classical and non-classical accounts of free will). The difference between **neo-classical compatibilism** and **classical compatibilism** has to do with how the free-will relatum of each view is fleshed out: the latter assumes that the classical account of free will is correct but the former does not.

Neo-Classical Incompatibilism: The view that is is metaphysically impossible for an ordinary human to exercise free will (where 'free will' is neutral between classical and non-classical accounts of free will) when **traditional determinism** is true because there is a type of antagonistic relevance relation between free will and determinism-related factors; alternatively: necessarily, determinism is *logically incompatible* with the **free-will thesis** (Mickelson [4]). The difference between neo-classical incompatibilism and **classical incompatibilism** is in the free-will relatum, namely that the latter presumes that the **classical account of free will** is correct and the former does not. Derk Pereboom's Four-Case Argument (Pereboom 2001, 2014) is a famous argument for neo-classical incompatibilism. See endnote 10 for further discussion.

Predeterminism: Predeterminism is a trans-temporal (past-to-future) form of determining (fixing, settling, etc.) of events and/or the truth-values of propositions, and its forms of determination include principles like bivalence, divine foreknowledge and providence, **traditional determinism**, eternalism, and socio-economic determinism.

Problem of Determinism: A loose collection of disagreements about how to spell out the relation and relata of the template-question "Does free will stand in relation R to determinism: *yes* or *no*?", and to explain how (if at all) asking and answering one or more instances of this template-question would help us to solve the **problem of free will**. While CAP theorists treated the **problem of determinism** narrowly as a problem of causal luck (i.e. the challenge of settling whether determinism-related causal and/or nomological factors preclude human free will), non-CAP theorists have suggested alternative characterizations (e.g. that determinism scenarios, like manipulation cases, sensitize us to threats posed by constitutive luck).

Problem of Free Will: The debate over the nature and existence of free will. In generic terms, the two basic solutions to the problem of free will are **free-willism** and **free-will denialism**. A complete solution to the problem of free will–and hence a complete statement of free-willism or free-will denialism–must spell out what free will is, e.g. by proposing a set of individually necessary and jointly sufficient conditions for acting freely.

Traditional Determinism: The doctrine that one unique future (relative to any arbitrary time t) is the inevitable result of the naturalistic factors which account for the evolution of the physical world over time (e.g. certain future-fixing causal and/or nomological relations between events in the past and events in the future). Notably, traditional determinism is not open to a so-called "broadly Humean" interpretation (Beebee and Mele 2002), for it affirms the presence of just the sort of necessity-in-nature that broadly Humean accounts of causation/laws of nature (by definition) reject; the doctrine known as "Humean determinism" is a species of **traditional indeterminism**. In this chapter, 'determinism' refers to traditional determinism and 'indeterminism' refers to traditional indeterminism unless otherwise stated. This is just one of many doctrines which goes by the name 'determinism.' (See also endnote 6 and White [3] for further discussion.)

Traditional indeterminism: The thesis that **traditional determinism** is false.

Notes

1 The term 'free-willist' has been used as an alternative name for the free-will libertarian (e.g. William James 1921, A.J. Ayer 1968, and Robert Kane 1996) and the term 'free-willism' is commonly associated with the theological position of Armenianism. We do not follow such usage here. The term 'free-will skepticism' is sometimes used to refer to denialism and/or an epistemic position about what we are justified in believing vis-a-vis the truth of denialism (e.g. McKenna and Pereboom 2016, p. 32); we editors prefer to restrict 'denialism' to a claim about the existence of free will and to restrict the term 'skepticism' to epistemic positions (e.g. the view that *we are justified in believing* that denialism is true and/or the more modest view that *we are not justified in believing* that free-willism is true).

2 No name was assigned to someone who embraced the conjunction of denialism and indeterminism, i.e. someone who answered the two questions above "no" and "no" (though 'hard indeterminist' would be apt.)

3 The terms 'compatibilism' and 'incompatibilism' were, by all appearances, coined by Keith Lehrer in the 1960s and were greatly popularized by Peter van Inwagen, especially van Inwagen 1983.

4 A *broadly libertarian account* of free will is one which proposes that it is at least metaphysically possible for someone to act freely, but includes at least one necessary condition which is metaphysically impossible to satisfy when determinism is true (e.g. Vicens [5], Adams [7], Smilansky [12], Balaguer [19]). Notably, one may adopt a broadly libertarian account of free will without accepting free-will libertarianism or anthropocentric possibilism (e.g. Pereboom 2014; Mickelson "Hard Times for Hard Incompatibilism", ms.).

5 Another notable feature of CAP is that CAP theorists typically focused on logical relationships between propositions rather than metaphysical relationships between non-propositional phenomena. For example, van Inwagen introduced a logical entailment thesis to capture the traditional metaphysical doctrine of determinism (see van Inwagen 1990, p. 277 for helpful diagrams), and used this entailment thesis as a proxy for traditional determinism in logic-text proofs which aimed to demonstrate that a *strict logical inconsistency* relation holds between between determinism and the **classical free-will thesis** (e.g. van Inwagen 1983; see also Mickelson [4]).

6 The traditional doctrine of determinism is interesting, in part, because it provides the limiting-case doctrine for minimal actual-sequence leeway, i.e. it states that there is literally zero indeterministic leeway (of any kind) in the world. Assuming determinism, not even God could intervene to prevent the "determined" future from coming to pass (e.g. van Inwagen 1990, p. 277, Sehon 2011, Mickelson 2012). As such, determinism was a useful tool for exploring free will as it was classically characterized. However, philosophers have provided interesting reasons for thinking that there are other–at least equally good or better–ways of defining 'determinism' vis-à-vis the problem of free will (e.g. Dennett 2003, White [3]).

7 The neo-classical compatibility problem is evident when philosophers frame the problem of determinism as a debate about whether determinism (determinism-related factors, deterministic causal luck, or the like) is a threat to free will because it keeps people from being able to act otherwise and/or because it keeps people from being an adequate source of their own actions (Mickelson [4]). When philosophers adopt this neo-classical framework, they struggle to classify views–such as **constitutive-luck source impossibilism (**discussed below)–which insist that it is impossible to act freely when determinism is true even though determinism itself is entirely irrelevant to free will. This is notable, given that worries about constitutive luck have been present in discussions of the problem of determinism since its inception, as can be seen in the surviving records of the debates between the Stoics and their critics (see Pereboom 2009, Ch. 2).

8 Semi-compatibilism is the result of another notable attempt to re-orient the problem of determinism around the specific type of control required for moral responsibility in order to evade the narrow use of 'free will' established by CAP theorists (e.g. Fischer 1994; Fischer and Ravizza 2006; Fischer and Ravizza 1998). While semi-compatibilists agree that determinism-related factors may undermine some types of control (e.g. "regulative control"), they insist that such factors do not undermine the control required for moral responsibility. The semi-compatibilist sidesteps a direct challenge to

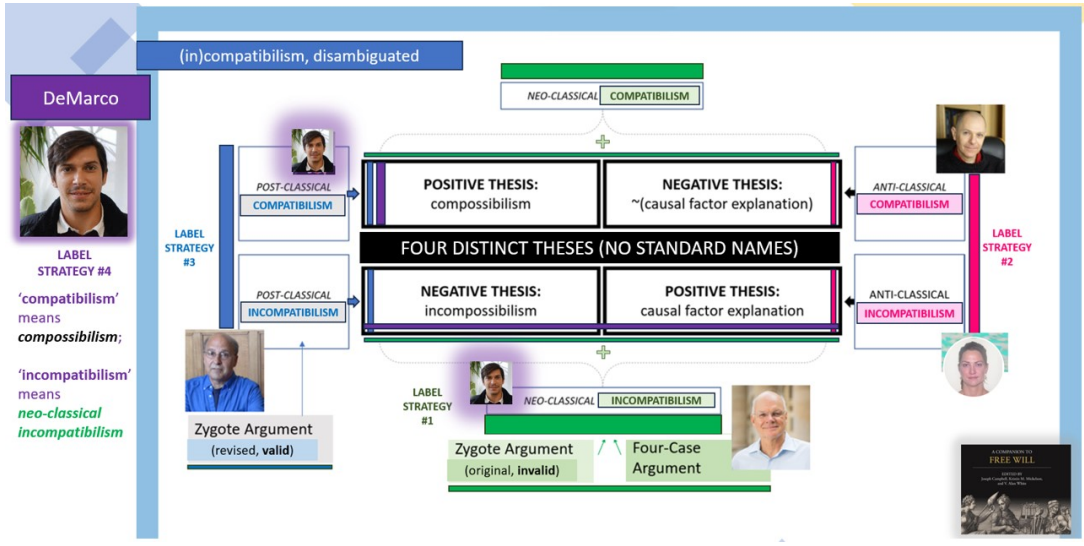
CAP theorists about the meaning of 'free will' by taking no stand on what this term does (or should) mean–except to say that *if* a person insists on using the classical definition of 'free will' and it turns out that determinism-related factors preclude one's *ability to do otherwise* (as classical incompatibilists claim), this result would not establish that determinism-related factors preclude moral responsibility; it would, rather, show that free will (i.e. the ability to do otherwise) is not required for moral responsibility. As such, semi-compatibilism is distinct from neo-classical forms of compatibilism which fix the referent of 'free will' to the control condition on moral responsibility only insofar as the latter take a stand on what free will is (and what 'free will' means) while the semi-compatibilists do not.

9 Although Pereboom is not an impossibilist, he is an anthropocentric impossibilist because he denies that a being who has the properties of an ordinary human (i.e. someone like us, as we are in the actual world) can satisfy the "law-overriding" source condition he forwards as part of his broadly libertarian account of free will. That is, hard incompatibilism is a type of anthropocentric impossibilism, but is not a type of impossibilism (for discussion, see Mickelson "Hard Times for Hard Incompatibilism", ms.).

10 As philosophers moved away from the original CAP-based definitions of 'compatibilism' and 'incompatibilism', they retrofitted the qualifier "classical" to these terms as a way of marking that departure. (The terms 'classical compatibilism' and 'classical incompatibilism' are also applied to views held by pre-CAP philosophers, which leads to complications we cannot address here.) Adding such qualifiers is one way to show due respect for the methodological principle that philosophers may define their jargon however they like while keeping tabs on the dialectically significant variations currently in use. Following suit, we have applied the qualifier "neo-classical" to single out the initial successors to the classical characterizations. Expanding this tracking device, we wish to identify a few additional recharacterizations which may be of interest to the reader.

As already noted, *neo-classical incompatibilism* has two defining tenets, namely incompossibilism and a positive explanatory thesis which states roughly that incompossibilism is true because determinism (determinism-related causal/nomological factors) deprive ordinary humans of free will; *neo-classical compatibilism* is also a two-tenet view, one tenet negates the negative thesis of neo-classical incompatibilism and the other negates its positive thesis (which means that neo-classical compatibilism is not equivalent to mere compossibilism (Mickelson 2012, 2015a)). Some philosophers have proposed that we use 'incompatibilism' to denote only the positive thesis of neo-classical incompatibilism and 'compatibilism' to name its negation (Levy 2011: p. 1, n. 1; Mickelson 2015b); let these be *anti-classical incompatibilism* and *anti-classical compatibilism*, respectively. Assuming this anti-classical revision, incompossibilism is not a defining tenet of incompatibilism but remains a corollary, so the anti-classical redefinition of 'incompatibilism' leaves the term's earlier neo-classical meaning largely intact. However, anyone who rejects anti-classical incompatibilism qualifies as an anti-classical compatibilist, which means that some *impossibilists* qualify as compatibilists on this anti-classical taxonomy. Anti-classical theorists consider this a feature rather than a bug, for it highlights that some philosophers argue for the negative thesis of neo-classical incompatibilism but against its positive thesis—a position that is not supposed to be available according to popular CAP-based narratives. (Kadri Vihvelin aims to achieve a similar goal via alternative terminological revisions, roughly: keep compossibilism as a defining tenet of 'compatibilism', redefine 'incompatibilism' to pick out the conjunction of incompossibilism and anthropocentric possibilism, and add 'impossibilism' to refer to anthropocentric impossibilism (e.g. Vihvelin 2008, 2013). A downside of this "tripartite taxonomy" is that anthropocentric impossibilists (e.g. hard incompatibilists) cannot be classified as incompatibilists even when they embrace both tenets of neo-classical incompatibilism (e.g. Vihvelin 2013: p.242, n. 5; Mickelson 2015a)). Other philosophers now use 'compatibilism' and 'incompatibilism' as their preferred labels for compossibilism and incompossibilism (e.g. Mele [30] and 2017: p. 6, n. 4; Capes [20]); to track this usage, let these be *post-classical compatibilism* and *post-classical incompatibilism*, respectively. These post-classical revisions bring back a bipartite taxonomy of (in)compatibilism by rejecting—fruitfully, according to post-classical theorists—more complicated taxonomies which treat the traditional dispute between anti-classical compatibilists and anti-classical incompatibilists as a fundamental point of divide in the contemporary free-will debate. A purported upside of the post-classical tax-

INFOGRAPHIC NOT IN ORIGINAL DOCUMENT



onomy is that impossibilists cannot be compatibilists (since post-classical compatibilists are compossibilists); a downside is that anti-classical compatibilists and anti-classical incompatibilists are lumped into one motley "anti-compatibilist" category. Hybrid recharacterizations are also found in the literature, e.g. using 'incompatibilism' to denote neo-classical incompatibilism but 'compatibilism' to denote mere compossibilism (see Mickelson 2021 for discussion). Despite appearances, the latter hybrid does not yield a genuine bipartite taxonomy, for (assuming these hybrid definitions) it may be that compatibilism and incompatibilism are both false and some third view—unnamed by the hybrid theorist—is true (e.g. constitutive-luck impossibilism).

With the above distinctions in hand, readers are better prepared to spot the common practice of technically defining 'incompatibilism' in one way while using it in another (e.g. McKenna and Pereboom 2016: pp. 30 and 151; Sartorio 2016: pp. 147 and 157) and to track fundamental differences between famous "arguments for incompatibilism". For example, Pereboom's Four-Case Argument (Pereboom 2001, 2014) concludes to neo-classical incompatibilism (Pereboom [29, n.5]), relying upon a slippery-slope argument to support post-classical incompatibilism and a modest best-explanation argument to support anti-classical incompatibilism; Alfred Mele's original Zygote Argument (Mele 2006) is invalid because its premises support mere post-classical incompatibilism but its conclusion is a statement of anti-classical incompatibilism (and/or neo-classical incompatibilism) (Mickelson 2015b); Mele's revised Zygote Argument (e.g. Mele 2013, 2017, 2019) is an argument for post-classical incompatibilism (a.k.a. incompossibilism) but it is not an argument for anti-classical incompatibilism (and hence is not an argument for neo-classical incompatibilism); Kristin Mickelson's Master Manipulation Argument—like Galen Strawson's Basic Argument (Strawson 1994)—concludes to impossibilism via reasoning which implies that post-classical incompatibilism (a.k.a. incompossibilism) is a true but metaphysically trivial position and that anti-classical incompatibilism (hence neo-classical incompatibilism) is false (e.g. Mickelson 2015b, 2019a, 2019b, "Hard Times for Hard Incompatibilism," ms.). Again, the novel qualifiers we have applied to the term 'incompatibilism' here are merely a rhetorical device for tracking the different views currently called by name 'incompatibilism'; whether philosophers should continue to use a single term in such disparate ways is another matter. Indeed, Mickelson argues that non-CAP theorists should sidestep the project of rehabilitating jargon from a research paradigm they reject and instead embrace new ways of talking about the fundamental divides in the contemporary debate (e.g. Mickelson [4]). Readers are advised to keep such differences and debates in mind as they read the chapters in this volume, and are invited to consider their own preferred solution to these jargon/taxonomy problems.

11 Since all the strands are critical of libertarianism, a CAP theorist may interpret these arguments as lending support to classical compatibilism.

12 Notably, other interesting problems arise when we untether the Consequence Argument from CAP, e.g. it is unclear that the argument still pinpoints determinism as a threat to free will (e.g. Campbell 2007).

13 Van Inwagen finds some logical space in the possibility of imminent or agent causation but this just raises further puzzles (van Inwagen 1983: 151–52).

14 Notably, this means that if the conclusion of Mickelson's Master Manipulation Argument is true, Pereboom's hard incompatibilism is also false. Readers should note that interesting dialectical points like this one are often obscured by the common classically-driven practice of lumping all manipulation arguments together under the label "arguments for incompatibilism." We advise the reader to look carefully at the stated conclusion of any given manipulation argument in order decide whether the argument aims to support mere incompossibilism, a type of classical or neo-classical incompatibilism, or a type of impossibilism which entails that incompossibilism is true but classical and neo-classical forms of incompatibilism are false.

15 If, as we claim, the background assumptions of much X-Phi inquiry is within the tradition of CAP, then many issues, such as the role that non-causal types of luck (e.g. constitutive luck) may play in deterministic or indeterministic scenarios, are completely left out of the picture, and this may well skew the subjects' responses in errant ways. Perhaps X-Phi studies might better reflect non-CAP concerns in the future? (For positive signs of movement in that direction, see Cova 2022.)

Bibliography

Ayer, A.J. (1968). *The Origins of Pragmatism*. London: Macmillan.

Beebee, H. and Mele, A. (2002). Humean compatibilism. *Mind* 111: 201–223.

Bobzien, S. (1998). *Determinism and Freedom in Stoic Philosophy*. Oxford: Oxford University Press.

Bobzien, S. (2021). *Determinism, Freedom, & Moral Responsibility: Essays in Ancient Philosophy*. Oxford: Oxford University Press.

Boolos, G. et al. (1990). An open peer commentary on the emperor's new mind. *Behavioral and Brain Sciences* 13 (4): 655.

Buckareff, A.A. (2019). Time, leeway, and the laws of nature: why humean compatibilists cannot be eternalists. *Metaphysica* 20 (1): 51–71.

Campbell, J.K. (2007). Free will and the necessity of the past. *Analysis* 67 (2): 105–111.

Clarke, R. (2003). *Libertarian Accounts of Free Will*. Oxford: Oxford University Press.

Cova, F. (2022). It was all a cruel angel's thesis from the start: folk intuitions about zygote cases do not support the zygote argument. In: *Advances in Experimental Philosophy of Free Will and Responsibility* (ed. T. Nadelhoffer and A. Monroe). London: Bloomsbury Academic.

Cova, F. et al. (2018). Estimating the reproducibility of experimental philosophy. *Review of Philosophy and Psychology* 12: 9–44.

Dennett, D. (2003). *Freedom Evolves*. London: Penguin Books.

Fischer, J.M. (1994). *The Metaphysics of Free Will: An Essay on Control*. Oxford: Blackwell.

Fischer, J.M. and Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility*. Cambridge: Cambridge University Press.

Fischer, J.M. and Ravizza, M. (2006). *My Way: Essays on Moral Responsibility*. New York: Oxford University Press.

Frankfurt, H.G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy* 66 (23): 829–839.

Kane, R. (2003). Free will: new directions for an ancient problem. In: *Blackwell Readings In Philosophy: Free Will*, 3e. Oxford: Wiley-Blackwell. 222-248.

Kane, R. (1996). *The Significance of Free Will*. New York: Oxford University Press.

Lehrer, K. (1990). *Metamind*. Oxford: Clarendon Press.

Levy, N. (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. New York: Oxford University Press.

Libet, B. et al. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential): the unconscious initiation of a freely voluntary act. *Brain* 106 (3): 623–642.

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

McKenna, M. and Pereboom, D. (2016). *Free Will: A Contemporary Introduction*. New York: Routledge.

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

Mele, A. (2017). *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*. Oxford: Oxford University Press.

Mele, A. (2019). *Manipulated Agents: A Window to Moral Responsibility*. Oxford: Oxford University Press.

Mickelson, K.M. (a.k.a. K. Demetriou). (2012). *Free Will Fundamentals: Agency, Determinism, and (In) compatibility.* [Dissertation, University of Colorado, Boulder]. https://scholar.colorado.edu/concern/graduate_thesis_or_dissertations/g732d9110.

Mickelson, K.M. (a.k.a. K. Demetriou). (2015a). A critique of Vihvelin's three-fold classification. *Canadian Journal of Philosophy* 45 (1): 85–99.

Mickelson, K.M. (a.k.a. K. Demetriou). (2015b). The zygote argument is invalid—now what? *Philosophical Studies* 172 (11): 2911–2929.

Mickelson, K.M. (a.k.a. K. Demetriou). (2017). The manipulation argument. In: *The Routledge Companion to Free Will* (ed. K. Timpe, M. Griffith and N. Levy). New York: Routledge. 166–178

Mickelson, K.M. (a.k.a. K. Demetriou). (2019a). Free will, self-creation, and the paradox of moral luck. *Midwest Studies in Philosophy* 43 (1): 224–256.

Mickelson, K.M. (a.k.a. K. Demetriou) (2019b). The problem of free will and determinism: an abductive approach. *Social Philosophy & Policy* 36 (1): 154–172.

Mickelson, K.M. (a.k.a. K. Demetriou). (2021). The zygote argument is still invalid: so what? *Philosophia* 49 (2): 705–722.

Nagel, T. (1976). Moral luck II. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 137–151.

Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.

Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. (2005). Surveying freedom: folk intuitions about free will and moral responsibility. *Philosophical Psychology* 18 (5): 561–584.

Nichols, S. and Knobe, J. (2007). Moral responsibility and determinism: the cognitive science of folk intuitions. *Noûs* 41 (4): 663–685.

Penrose, R. (1989). *The Emperor's New Mind*. Oxford: Oxford University Press.

Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.

Pereboom, D. (2009). *Free Will*, 2e (ed. D. Pereboom). Indianapolis: Hackett Publishing Company.

Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. New York: Oxford University Press.

Sartorio, C. (2016). *Causation and Free Will*. Oxford: Oxford University Press.

Sehon, S. (2010). A flawed conception of determinism in the consequence argument. *Analysis* 71: 30–38.

Shabo, S. (2011). "What must a proof of incompatibilism prove?" *Philosophical Studies* 154: 361–371.

Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies* 75 (1/2): 5–24.

Strawson, G. (2011). Free will. In: *Routledge Encyclopedia of Philosophy* (ed. E. Craig). London: Routledge. https://www.rep.routledge.com/articles/thematic/free-will/v-1/sections/pessimism.

Strawson, P. (1962). Freedom and Resentment. *Proceedings of the British Academy* 48: 1–25.

Swenson, P. (2016). Ability, foreknowledge, and explanatory dependence. *Australasian Journal of Philosophy* 94: 658–71.

Taylor, R. (1963). *Metaphysics*, 1e. Englewood Cliffs: Prentice-Hall.

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.

van Inwagen, P. (1998). The mystery of metaphysical freedom. In: *Metaphysics: The Big Questions* (ed. P. van Inwagen and D.W. Zimmerman). Oxford: Blackwell Readings in Philosophy. 365–373.

van Inwagen, P. (2000). Free will remains a mystery. *Philosophical Perspectives* 14: 1–19.

van Inwagen, P. (2017). The problem of fr** w*ll. In: *Thinking about Free Will*, Cambridge: Cambridge University Press. 192–209.

Vihvelin, K. (2008). Compatibilism, incompatibilism, and impossibilism. In: *Contemporary Debates in Metaphysics* (ed. J. Hawthorne, T. Sider and D. Zimmerman). Malden: Blackwell. 303–318.

Vihvelin, K. (2013). *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. New York: Oxford University Press.

Vilhauer, B. (2012). Taking free will skepticism seriously. *Philosophical Quarterly* 62: 833–852.

Williams, B. and Nagel, T. (1976). Moral luck. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 115–135 + 137–151.

Part I

Preliminaries

4

(In)compatibilism

KRISTIN M. MICKELSON

1 The Problem of Determinism and Its Candidate Solutions

The terms 'compatibilism' and 'incompatibilism' were coined in the early 1960s and quickly became part of the basic vocabulary of *the classical analytic paradigm*, the dominant research paradigm of the classical period in the free-will debate (c. 1965–1985).[1] Philosophers working in the classical analytic paradigm, hereafter *classical analytic theorists*, used the phrase 'free-will thesis' to name the thesis that we, that is, ordinary human beings, exercise free will. The term 'incompatibilism' was introduced to name the view that the free-will thesis is *incompatible* with determinism, and 'compatibilism' was introduced to name the view that the free-will thesis is *compatible* with determinism.[2] Following Peter van Inwagen – who crystallized the distinctive assumptions and methods that structured the classical analytic paradigm in *An Essay on Free Will* (1983) – most philosophers still speak as if the central problem of free will and determinism is *The Compatibility Problem*, roughly the challenge of settling whether compatibilism or incompatibilism is true.

The classical analytic paradigm started to degenerate almost as soon as it formed, with the advent of "Frankfurt examples" (see Haji Ch. 6 this volume) playing a critical role in its demise. As one would expect, framing the free-will debate in the anachronistic jargon of a degenerated research paradigm is not a benign practice. In this section, I shed new light on the structure of the classical analytic Compatibility Problem by juxtaposing it with a more contemporary – and, at the same time, more traditional – way of characterizing the problem of free will and determinism and its array of candidate solutions. In the next section, I look more closely at the correlation and relevance relations which classical analytic theorists – and those working in their wake – have conflated under the label "incompatibility." In doing so, I clear the way for contemporary free-will theorists to escape the jargon, narratives, and question-begging background assumptions of the outdated classical analytic paradigm.

A Companion to Free Will, First Edition. Edited by Joseph Campbell, Kristin M. Mickelson, and V. Alan White.

4

(In)compatibilism

KRISTIN M. MICKELSON

1 The Problem of Determinism and Its Candidate Solutions

The terms 'compatibilism' and 'incompatibilism' were coined in the early 1960s and quickly became part of the basic vocabulary of *the classical analytic paradigm*, the dominant research paradigm of the classical period in the free-will debate (c. 1965–1985).[1] Philosophers working in the classical analytic paradigm, hereafter *classical analytic theorists*, used the phrase 'free-will thesis' to name the thesis that we, that is, ordinary human beings, exercise free will. The term 'incompatibilism' was introduced to name the view that the free-will thesis is *incompatible* with determinism, and 'compatibilism' was introduced to name the view that the free-will thesis is *compatible* with determinism.[2] Following Peter van Inwagen – who crystallized the distinctive assumptions and methods that structured the classical analytic paradigm in *An Essay on Free Will* (1983) – most philosophers still speak as if the central problem of free will and determinism is *The Compatibility Problem*, roughly the challenge of settling whether compatibilism or incompatibilism is true.

The classical analytic paradigm started to degenerate almost as soon as it formed, with the advent of "Frankfurt examples" (see Haji Ch. 6 this volume) playing a critical role in its demise. As one would expect, framing the free-will debate in the anachronistic jargon of a degenerated research paradigm is not a benign practice. In this section, I shed new light on the structure of the classical analytic Compatibility Problem by juxtaposing it with a more contemporary – and, at the same time, more traditional – way of characterizing the problem of free will and determinism and its array of candidate solutions. In the next section, I look more closely at the correlation and relevance relations which classical analytic theorists – and those working in their wake – have conflated under the label "incompatibility." In doing so, I clear the way for contemporary free-will theorists to escape the jargon, narratives, and question-begging background assumptions of the outdated classical analytic paradigm.

A Companion to Free Will, First Edition. Edited by Joseph Campbell, Kristin M. Mickelson, and V. Alan White.

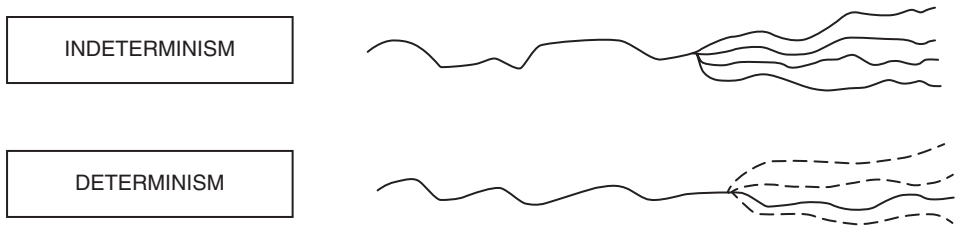
1.1 Defining Determinism

Let us start with the notion of determinism that is central to the traditional problem of free will and determinism (hereafter, *the problem of determinism*). William James, who was one of the first philosophers to use the term 'determinism' in the free-will literature, characterized the doctrine as follows:

> What does determinism profess? It professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be. The future has no ambiguous possibilities hidden in its womb; the part we call the present is compatible with only one totality. Any other future complement than the one fixed from eternity is impossible. The whole is in each and every part, and welds it with the rest into an absolute unity, an iron block, in which there can be no equivocation or shadow of turning. (James 1884/1897)

In this passage, James describes a universe in which one unique future is predetermined (unconditionally pre-fixed, made inevitable) by factors which were fully in place before that future unfolds.[3]

We can emphasize the radical nature of determinism by contrasting it with the type of "garden of forking paths" timeline that unfolds when determinism is false (i.e. when indeterminism is true) with the following pair of diagrams from Peter van Inwagen (1990):



In the top diagram, the solid lines represent that there is at least one point in this timeline relative to which there are several open "alternative futures." A person in the top timeline would have access to multiple futures, analogous to the way in which a person has access to different routes forward when she comes to a fork in a river or road (van Inwagen 1990, p. 277). By contrast, the bottom timeline shows that when determinism is true, we never "confront a sheath of possible futures" in the actual timeline of our lives (van Inwagen 1990, p. 277); only one timeline is actually open to us, and every apparent fork in the road is an illusion.

The idea that the future is made inevitable by naturalistic forces beyond our control has been driving the free-will debate for millennia, having its roots in the naturalistic (cause-and-effect) account of fate developed by the Stoics (Bobzien 1998; Pereboom 2009, pp. 5–16; see also Kane 2002, p. 6).[4] For present purposes, I will set aside questions about precisely which features of the world do the critical future-fixing work when determinism is true (e.g. causation, laws of nature, etc.). For simplicity, I will use the phrase "determinism-related factors" to refer to those features of the world (whatever they are) which account for the state-by-state and moment-by-moment evolution of the world when determinism is true, that is, the factors which do the "work" of making one unique future inevitable when they obtain.

In my estimation, the traditional debate over the *relationship* between free will and determinism rightly captures our attention because it forces us to confront *the problem of free will*. In short, the problem of free will is the challenge of identifying the *nature* of free will so that

we can answering the more practically pressing question of whether or not we (ordinary humans) *have* free will. When the dialectical role of determinism is seen in this way – that is, as a way of generating and evaluating candidate solutions to the problem of free will – it appears that the problem of determinism has two main components, what I call *the correlation problem* and *the explanation problem* (Mickelson 2019a, 2019b, 2021a).[5] Let us consider these two problems in turn.

1.2 The Correlation Problem

The central challenge of the correlation is problem is captured in one familiar question:

> **The Correlation Problem**: Is it metaphysically possible for a typical human to exercise free will in a world at which determinism is true?[6]

The long history of the free-will debate demonstrates that people deeply disagree about whether the answer to this question is "yes" or "no." Those who think that the correct answer to the correlation problem is "no" thereby embrace the view that free will – whatever else might be true of it – is not the sort of thing that could be exercised by a normal human in a world at which determinism is true. For these theorists, the intuitive "no free will" judgment elicited by thought experiments involving humans acting in a world at which determinism is true is a *data point* that must be accommodated by any viable theory of free will. That is, to answer "no" to the correlation question is to commit oneself to a desideratum on any viable theory of free will, namely: it must include at least one necessary condition which cannot be satisfied by a normal human living in a world at which determinism is true. By contrast, those who answer "yes" thereby reject that there is any such desideratum on a viable account of free will. In other words, determinism scenarios reveal that people not only have different intuitions about what free will is, but they also disagree about the standards by which a theory of free will is to be judged.

Despite the dialectical significance of the "yes" and "no" answers to the correlation problem, there is no standard jargon which unequivocally tracks the modal views associated with them. To remedy this, I have introduced new terminology (Mickelson 2012, 2015a, 2015b, 2017, 2019a, 2019b, 2021a). To give a "no" response to the correlation problem is to claim that human free will and determinism-related factors are *incompossible*, that is, it is metaphysically impossible for them to co-exist. In light of this, I say that those who answer "no" to the correlation problem are proponents of an *incompossibility solution* to the correlation problem, while those who answer "yes" endorse a *compossibility solution*. Correspondingly, *incompossibilism* is the view that the incompossibility solution to the correlation problem is correct, and *compossibilism* is the view that the compossibility solution is correct.

2 The Explanation Problem

Incompossibilism answers one problem only to raise another: unlike the compossibilist, the incompossibilist must explain or account for the incompossibility of human freedom and determinism-related factors. To solve this *explanation problem*, the incompossibilist must answer the following questions:

> **The Unmet Condition Challenge** (**E1**): Which necessary condition C on free will cannot be satisfied by a normal human when determinism is true?

The Condition Underminer Challenge (**E2**): What prevents a normal human from satisfying condition C when determinism is true?[7]

Although the E1/E2 distinction is present in contemporary discussions of free will, free-will specialists have not made an effort to track this distinction or that between an incompossibility solution to the correlation problem and extant solutions to E2 of the explanation problem. Presumably, this is at least partly because their understanding of the problem of determinism has been shaped and constrained by the distinctive jargon and narratives of the classical analytic paradigm. Given these historical considerations, let us begin by looking at the bipartite explanation problem through the narrow lens of the classical analytic paradigm.

2.1 Classical Analytic Solutions to the Correlation and Explanation Problems

Within the classical analytic paradigm, the questions associated with E1 and E2 were not distinguished from a "no" answer to the correlation problem, and it's easy to understand why. Given the background assumptions of that paradigm, a complete solution to the explanation problem follows directly from an incompossibility solution to the correlation problem. How so? Classical analytic theorists agreed that free will is (or requires) an *ability to do otherwise*. Since it is standard practice to attach the term 'classical' to view which presuppose that free will is an ability to do otherwise, it is natural to refer to the incompossibilist who accepts the classical view of free will as a *classical incompossibilist*. The classical incompossibilist must interpret her solution to the correlation problem as providing key information about what this freedom-relevant ability to do otherwise amounts to. Specifically, the classical incompossibilist commits to the view that free will, that is, the ability to do otherwise, cannot be exercised in which one unique future is made inevitable by determinism-related factors (given the facts of the past). Free will is a more robust ability to do otherwise, roughly an ability to "settle" which of the multiple futures available to her at the moment of action, in the forking-paths timeline of her life, comes to pass; in counterfactual terms, an actor exercises the relevant ability to do otherwise in performing A only when it is true of the actor that, holding fixed the laws and the facts of the past prior to the action, the actor still *could have done otherwise* than perform A, that is, she could have *avoided* or *refrained* from doing what she actually did and done something else instead (e.g. van Inwagen 1983, 2017; Balaguer Ch. 9 this volume).

Notably, classical analytic theorists also took for granted the truth of *anthropocentric possibilism*, the view that it is at least metaphysically possible for an ordinary human to exercise free will (e.g. van Inwagen 1983; Clarke 2003; Vihvelin 2013).[8] Any account of free will that is consistent with both incompossibilism and anthropocentric possibilism is now commonly classified as a *broadly libertarian account* of free will.[9] This means that all incompossibilists working in the classical analytic paradigm were committed to a broadly libertarian solution to E1.[10] I will hereafter refer to the general solution to E1 endorsed by classical incompossibilists, according to which some libertarian interpretation of the ability to do otherwise is correct, the *classical libertarian solution* to E1.[11]

According to the classical libertarian solution to E1, it is metaphysically possible for a normal human to exercise free will only if there is some type of *actual-sequence leeway* in the evolution of the universe. There is actual-sequence leeway in the evolution of the universe only if there is at least one point in time at which, holding fixed the laws and facts of the past, more than one future might unfold. That is, more than one future is *accessible* from some point in the actual timeline, just as we see in van Inwagen's "forking paths" depiction of

indeterminism (in van Inwagen's diagram seen earlier). Recall that determinism, by definition, is true only when there is literally zero actual-sequence leeway in the evolution of the physical world, that is, when exactly zero "forking paths" are permitted in the actual timeline. This means that if determinism were true, certain determinism-related factors would rule out the sort of actual sequence leeway – *whatever type that may be* – that an exercise of free will requires (given the libertarian take on the notion of an ability to do otherwise). Within the classical analytic paradigm, then, an incompossibility solution to the correlation problem implies the classical libertarian solution to E1, which in turn implies the *classical incompatibilist solution* to E2: when determinism is true, it is *owing to determinism-related factors* that people cannot exercise the ability to do otherwise that an act of free will requires. For ease of reference, I will hereafter refer to this line of reasoning – from incompossibilism, via the classical libertarian solution to E1, to the classical incompatibilist solution to E2 – as the *classical bridge inference*.[12]

Now that we have a better grasp on the classical analytic characterization of the problem of determinism, we can understand why classical analytic theorists believed – as their Compatibility Problem narrative tells us – that the problem admits only two (mutually exclusive and collectively exhaustive) candidate solutions. On the one hand, classical analytic theorists could adopt an incompossibility solution to the correlation problem which, given the background assumptions of the classical analytic paradigm, came theoretically loaded – via the classical bridge inference – with the classical libertarian solution to E1 and the classical incompatibilist solution to E2. This complex position is now known as *classical incompatibilism*. On the other hand, the classical analytic theorists could embrace a compossibility solution to the correlation problem and proclaim that the explanation problem is a pseudo-problem generated by a false (incompossibility) solution to the correlation problem and, so, all candidate solutions to the explanation problem are wrong. The latter position is now known as *classical compatibilism*. For those working within the classical analytic paradigm, then, it was quite sensible to summarize the problem of determinism as a debate about whether classical compatibilism or classical incompatibilism is true.

2.2 Contemporary Solutions to E1 of the Explanation Problem

The classical analytic paradigm has fallen out of favor, largely because many philosophers are unwilling to grant its presumptions that free will is/requires an ability to do otherwise and that anthropocentric possibilism is true.[13] Indeed, a growing number of contemporary philosophers explicitly reject both anthropocentric possibilism and the classical definition of 'free will'. For example, Galen Strawson (1986, 1994, 2002, 2008, 2011) and Derk Pereboom (2001, 2014) are incompossibilists who reject the classical libertarian solution to E1, favoring instead a type of *source solution* to E1 according to which free will requires one to be the genuine *source* of one's action. Although their accounts of freedom-relevant sourcehood are substantively different, Pereboom and G. Strawson agree that it is metaphysically impossible for a normal human to satisfy the necessary source condition on free will.[14] We might say, then, that G. Strawson and Pereboom are each *source incompossibilists*, insofar as each embraces a strict source solution to E1 and, moreover, each is an *anthropocentric source impossibilist*, insofar as each embraces the view that it is metaphysically impossible for an ordinary human to satisfy the source condition on free will. However, whereas G. Strawson's understanding of the source condition leads him to endorse unqualified *impossibilism*, that is, the unqualified view that it is metaphysically impossible for any being to exercise free will, Pereboom's understanding of the source condition leads him to reject impossibilism (e.g. Pereboom 2001, p. 132).

In addition, some philosophers have suggested a type of hybrid solution to E1, proposing that an ability to do otherwise is required to satisfy the source condition on free will. For example, Robert Kane's "self-forming actions" account of free will seems to be a hybrid solution of this kind (e.g. Kane 1996, 2004, 2016, 2019; see also Balaguer CH. 9 this volume). As such, it would now be question-begging to frame the free-will debate, as classical analytic theorists once did, against the background assumptions that anthropocentric possibilism is true and that some version of the classical (ability-to-do-otherwise) characterization of free will is correct.

The fact that contemporary incompossibilists disagree about the nature of free will in a more fundamental way than their classical predecessors makes it easier to understand why contemporary free-will theorists struggle to provide a precise yet reasonably uncontroversial characterization of *free will* (and/or definition of 'free will') for use in their debate. The more precise one's proposed characterization, the more likely that one camp or another in the debate will find it question-begging; the more generic the proposal, the more likely that the referent of 'free will' will not be fixed securely enough to ensure that the interlocutors in the debate are disagreeing about the same thing (rather than having a mere verbal dispute in which two groups talk past each other because they are using the same phrase 'free will' to pick out different things).

There is no simple fix for this dialectical difficulty, but neo-classical and non-classical theorists have generally agreed to fix the referent of 'free will' by saying that free will is a type of *control* or *up-to-one-ness* that is necessary for moral responsibility, such that a person is morally responsible for an action A only if she exercises this type of control or exhibits this type of up-to-one-ness when performing A. Since it would be easy to find objections to any attempt to give a dialectically neutral characterization of free will (or definition of 'free will'), I will not attempt to provide one here; I leave readers to weigh the pros and cons of popular working definitions of 'free will' for themselves.

2.3 Contemporary Solutions to E2 of the Explanation Problem

Philosophers who are unaware of recent shifts away from the classical analytic paradigm may find it strange, even absurd, to suggest that there is room for a substantive and philosophically interesting debate about what *explains* and/or *accounts for* the lack of free agents in worlds at which determinism is true. Readers in this camp might wonder: Assuming that incompossibilism is true, isn't it just obvious – so obvious that it may go unstated – that certain *determinism-related factors* play at least some role in making people unfree when determinism is true? Does it really matter in the end whether the incompossibilists disagree about the fine-grained details regarding *which* determinism-related factors – the causation, the laws of nature, a conjunction of causal factors and facts about the past, etc. – pose a threat to human freedom, or whether determinism-related factors preclude free will *on their own* or only in conjunction with some other factors beyond our control?

Readers drawn to such questions must be reminded that the contemporary free will debate is no longer constrained by the assumptions and methods of the classical analytic paradigm. Once one abandons the classical analytic stipulation that 'free will' denotes an ability to do otherwise, the classical bridge inference is no longer a viable way to close the "explanatory gap" between an incompossibility solution to the correlation problem and the classical incompatibilist solution to E2 of the explanation problem. Indeed, for all that is stated by incompossibilism, it might simply be a *brute fact* that no one acts freely when determinism is true.[15] Of course, few would accept that the incompossibility of free will and determinism-related

factors is a brute fact; assuming incompossibilism is true, there must be some *better explanation* for its truth than that. But what *is* that better explanation? This brings us to the central point of this section: influential incompossibilists can and do disagree about what makes people unfree when determinism is true. That is, contemporary incompossibilists, unlike their classical predecessors, substantively disagree about the correct solution to E2.

The suggestion that the literature already includes a variety of philosophically interesting and conflicting solutions to E2 may take some readers by surprise. After all, philosophers generally speak as though all incompossibilists agree that determinism-related factors keep people from acting freely when determinism is true, but are divided about whether such factors preclude free will *because they keep people from exercising the ability to do otherwise* or, rather, such factors preclude free will *because they keep people from being the source of their own actions* (e.g. Fischer and Ravizza 1998, p. 151, Kane 2002, p. 6; Griffith 2017, p. 2; Vihvelin 2022). Unfortunately, the existence and contours of this lively in-house dispute among incompossibilists have been obscured in part because the dispute cannot be adequately characterized using the jargon and narratives of the classical analytic paradigm.

For purposes of categorizing extant solutions to E2, it is helpful to distinguish between two basic categories of factors which are, arguably, *beyond one's control*. First, there are two types of actor-extrinsic factors. On the one hand, there are states of physical world outside (i.e. external to the properties which constitute) a given actor. These actor-extrinsic factors include both aspects of one's immediate environment (e.g. the physical properties of the room one is sitting in) and complete states of the world in the remote past, for example, states which were obtained before the first human was born (e.g. there were no humans when dinosaurs roamed the Earth). On the other hand, there are actor-extrinsic factors which govern or otherwise account for the way in which the universe evolves from one state to another over time. For example, *determinism-related factors* are those factors which, given the state of the physical universe at one time, ensure the fine-grained state of the physical universe at the next moment--and every moment thereafter, until the end of time. The specific nature of the factors which account for the evolution of the universe is a matter of debate, for example, they may be causal relations, laws of nature, and so on. Second, there are features of an actor which may be (arguably, at least) beyond an actor's control. These include the properties which an actor has at a given moment (e.g. one's precise brain state at a given moment), features an actor has during their entire lifespan (e.g. one's genetic endowment), as well as an actor's essential properties, if there are such things (e.g. those properties which establish an actor's cross-world identity, such that we can posit that an actor in one possible world is – or at least has a counterpart in – some other possible world).

We might introduce a completely new set of labels to track the three mutually exclusive and collectively exhaustive classes of factors described earlier, but I will instead import and refine a few terms that have become commonplace in discussions of the paradox of moral luck (e.g. Hartman Ch. 10 this volume). This choice reflects my conviction that the problem of determinism and the paradox of moral luck are best understood as two rhetorically distinct frameworks for investigating *the problem of free will*, roughly the problem of identifying what free will is and whether we (i.e. ordinary humans) have it.[16] Like the problem of determinism, the paradox of moral luck is grounded partly in the observation that our normal practices of blame and punishment seemingly presuppose that we have the type of control required for basic-desert moral responsibility (McCormick Ch. 24 this volume), and yet close examination of individual actions indicates that no human has that type of control – rather, every human action is settled mostly – if not entirely – by factors beyond our control.[17] Moral-luck theorists have normalized the practice of using 'luck' as shorthand for the more cumbersome phrase

"factors beyond one's control" (Hartman 2017, pp. 23–31; Anderson 2019; Statman 2019), and I will follow that convention here. Inspired by Nagel (1976, 1979, 1986), I will speak of three basic types of luck: causal luck, circumstantial luck, and constitutive luck.[18]

For present purposes, let us say that a person is subject to *causal luck* when one lacks control over the factors which settle how the world evolves from one state of affairs to another (e.g. causal relations, causal laws, laws of nature, or the like). Let us say that one is subject to *circumstantial luck* when one lacks control over the non-causal/agent-extrinsic states of affairs (e.g. states of the universe prior to one's birth). Finally, one is subject to *constitutive luck* when one lacks control over one's own constitutive properties (e.g. one's genetic endowments with regards to intelligence and personality, or that one is an ordinary human rather than some other type of subhuman or superhuman creature).[19] Using these three categories of luck, we can easily identify a variety of extant solutions to E2: *causal luck solutions*, *circumstantial luck solutions*, *constitutive luck solutions*, and *hybrid solutions*. Let's consider each in turn.

A strict *causal luck solution* to E2 proposes that it is impossible for people to act freely when determinism is true because there are certain nomological and/or causation-related factors which make it the case that no human can exercise free will. The classical incompatibilist ("owing to determinism") solution to E2 may be classified as a type of causal luck solution. That said, there is plenty of room to wonder what this classical causal luck solution really amounts to. Philosophers commonly speak as though determinism per se precludes free will. Taken literally, however, this is an untenable response to E2. Why? Determinism has the ontological status of a proposition, and this proposition merely provides a *description* of a certain kind of world. There is little reason to suppose that a description (even if true) could itself pose a threat to free will; it is not the description, but *what it describes*, that has the ontological standing to preclude the exercise of free will (e.g. Hermes and Campbell 2012; Hermes 2013).

Philosophers who contend that certain evolution-governing factors *described by* determinism have the ontological standing required to undercut human freedom (as opposed to the thesis of determinism itself) carry the dialectical burden of specifying *which* evolutionary factors do the freedom-undermining work when determinism is true. Philosophers often pinpoint the *deterministic* quality of causation (laws of nature, or the like) which obtain when determinism is true as a specific threat to free will (e.g. Pereboom 2001, 2014; Sartorio 2016). Notably, though, the laws of nature may be deterministic in a perfectly standard sense of the term 'deterministic' even when the type of zero-leeway determinism described by James and depicted by van Inwagen (above) is false (e.g. Stone 1998; Dennett 2003; Sehon 2010; Mickelson 2012, 2019a, 2019b).[20] Other potentially freedom-relevant features of the laws of nature described by determinism are their *strength*, that is, that they are "strong" rather than "weak" (Perry 2004), and their *potency*, that is, that such laws are "unconditional" rather than "conditional" (Mickelson 2019a).[21] Given these distinctions, there is ample dialectical space for a lively in-house debate among the subset of incompossibilists who endorse a causal luck solution to E2 – a group we might call *causal luck incompossibilists* – to disagree about which determinism-related factors (if any) are antagonistic to free will. Given that some determinism-related factors are present even when determinism is false, such disagreements may have very interesting implications for the general project of assessing which account of free will is best (Mickelson 2019a and "Hard Times for Hard Incompatibilism," ms.).

Incompossibilists who reject broadly libertarian solutions to E1 may – and arguably must[22] – categorically reject causal luck solutions to E2. For example, G. Strawson is an impossibilist, and *a fortiori* an incompossibilist, who rejects the classical solutions to E1 and E2. With his Basic Argument, G. Strawson argues that we ("the folk") intuitively accept that there is an "ultimate starting point" source condition on free will, but the type of "buck stops here"

"factors beyond one's control" (Hartman 2017, pp. 23–31; Anderson 2019; Statman 2019), and I will follow that convention here. Inspired by Nagel (1976, 1979, 1986), I will speak of three basic types of luck: causal luck, circumstantial luck, and constitutive luck.[18]

For present purposes, let us say that a person is subject to *causal luck* when one lacks control over the factors which settle how the world evolves from one state of affairs to another (e.g. causal relations, causal laws, laws of nature, or the like). Let us say that one is subject to *circumstantial luck* when one lacks control over the non-causal/agent-extrinsic states of affairs (e.g. states of the universe prior to one's birth). Finally, one is subject to *constitutive luck* when one lacks control over one's own constitutive properties (e.g. one's genetic endowments with regards to intelligence and personality, or that one is an ordinary human rather than some other type of subhuman or superhuman creature).[19] Using these three categories of luck, we can easily identify a variety of extant solutions to E2: *causal luck solutions*, *circumstantial luck solutions*, *constitutive luck solutions*, and *hybrid solutions*. Let's consider each in turn.

A strict *causal luck solution* to E2 proposes that it is impossible for people to act freely when determinism is true because there are certain nomological and/or causation-related factors which make it the case that no human can exercise free will. The classical incompatibilist ("owing to determinism") solution to E2 may be classified as a type of causal luck solution. That said, there is plenty of room to wonder what this classical causal luck solution really amounts to. Philosophers commonly speak as though determinism per se precludes free will. Taken literally, however, this is an untenable response to E2. Why? Determinism has the ontological status of a proposition, and this proposition merely provides a *description* of a certain kind of world. There is little reason to suppose that a description (even if true) could itself pose a threat to free will; it is not the description, but *what it describes*, that has the ontological standing to preclude the exercise of free will (e.g. Hermes and Campbell 2012; Hermes 2013).

Philosophers who contend that certain evolution-governing factors *described by* determinism have the ontological standing required to undercut human freedom (as opposed to the thesis of determinism itself) carry the dialectical burden of specifying *which* evolutionary factors do the freedom-undermining work when determinism is true. Philosophers often pinpoint the *deterministic* quality of causation (laws of nature, or the like) which obtain when determinism is true as a specific threat to free will (e.g. Pereboom 2001, 2014; Sartorio 2016). Notably, though, the laws of nature may be deterministic in a perfectly standard sense of the term 'deterministic' even when the type of zero-leeway determinism described by James and depicted by van Inwagen (above) is false (e.g. Stone 1998; Dennett 2003; Sehon 2010; Mickelson 2012, 2019a, 2019b).[20] Other potentially freedom-relevant features of the laws of nature described by determinism are their *strength*, that is, that they are "strong" rather than "weak" (Perry 2004), and their *potency*, that is, that such laws are "unconditional" rather than "conditional" (Mickelson 2019a).[21] Given these distinctions, there is ample dialectical space for a lively in-house debate among the subset of incompossibilists who endorse a causal luck solution to E2 – a group we might call *causal luck incompossibilists* – to disagree about which determinism-related factors (if any) are antagonistic to free will. Given that some determinism-related factors are present even when determinism is false, such disagreements may have very interesting implications for the general project of assessing which account of free will is best (Mickelson 2019a and "Hard Times for Hard Incompatibilism," ms.).

Incompossibilists who reject broadly libertarian solutions to E1 may – and arguably must[22] – categorically reject causal luck solutions to E2. For example, G. Strawson is an impossibilist, and *a fortiori* an incompossibilist, who rejects the classical solutions to E1 and E2. With his Basic Argument, G. Strawson argues that we ("the folk") intuitively accept that there is an "ultimate starting point" source condition on free will, but the type of "buck stops here"

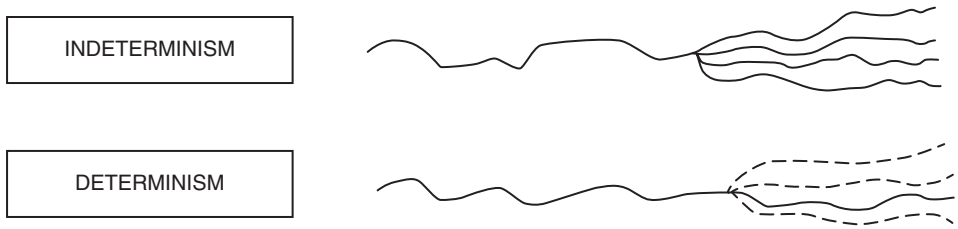
1.1 Defining Determinism

Let us start with the notion of determinism that is central to the traditional problem of free will and determinism (hereafter, *the problem of determinism*). William James, who was one of the first philosophers to use the term 'determinism' in the free-will literature, characterized the doctrine as follows:

> What does determinism profess? It professes that those parts of the universe already laid down absolutely appoint and decree what the other parts shall be. The future has no ambiguous possibilities hidden in its womb; the part we call the present is compatible with only one totality. Any other future complement than the one fixed from eternity is impossible. The whole is in each and every part, and welds it with the rest into an absolute unity, an iron block, in which there can be no equivocation or shadow of turning. (James 1884/1897)

In this passage, James describes a universe in which one unique future is predetermined (unconditionally pre-fixed, made inevitable) by factors which were fully in place before that future unfolds.[3]

We can emphasize the radical nature of determinism by contrasting it with the type of "garden of forking paths" timeline that unfolds when determinism is false (i.e. when indeterminism is true) with the following pair of diagrams from Peter van Inwagen (1990):



In the top diagram, the solid lines represent that there is at least one point in this timeline relative to which there are several open "alternative futures." A person in the top timeline would have access to multiple futures, analogous to the way in which a person has access to different routes forward when she comes to a fork in a river or road (van Inwagen 1990, p. 277). By contrast, the bottom timeline shows that when determinism is true, we never "confront a sheath of possible futures" in the actual timeline of our lives (van Inwagen 1990, p. 277); only one timeline is actually open to us, and every apparent fork in the road is an illusion.

The idea that the future is made inevitable by naturalistic forces beyond our control has been driving the free-will debate for millennia, having its roots in the naturalistic (cause-and-effect) account of fate developed by the Stoics (Bobzien 1998; Pereboom 2009, pp. 5–16; see also Kane 2002, p. 6).[4] For present purposes, I will set aside questions about precisely which features of the world do the critical future-fixing work when determinism is true (e.g. causation, laws of nature, etc.). For simplicity, I will use the phrase "determinism-related factors" to refer to those features of the world (whatever they are) which account for the state-by-state and moment-by-moment evolution of the world when determinism is true, that is, the factors which do the "work" of making one unique future inevitable when they obtain.

In my estimation, the traditional debate over the *relationship* between free will and determinism rightly captures our attention because it forces us to confront *the problem of free will*. In short, the problem of free will is the challenge of identifying the *nature* of free will so that

we can answering the more practically pressing question of whether or not we (ordinary humans) *have* free will. When the dialectical role of determinism is seen in this way – that is, as a way of generating and evaluating candidate solutions to the problem of free will – it appears that the problem of determinism has two main components, what I call *the correlation problem* and *the explanation problem* (Mickelson 2019a, 2019b, 2021a).[5] Let us consider these two problems in turn.

1.2 The Correlation Problem

The central challenge of the correlation is problem is captured in one familiar question:

> **The Correlation Problem**: Is it metaphysically possible for a typical human to exercise free will in a world at which determinism is true?[6]

The long history of the free-will debate demonstrates that people deeply disagree about whether the answer to this question is "yes" or "no." Those who think that the correct answer to the correlation problem is "no" thereby embrace the view that free will – whatever else might be true of it – is not the sort of thing that could be exercised by a normal human in a world at which determinism is true. For these theorists, the intuitive "no free will" judgment elicited by thought experiments involving humans acting in a world at which determinism is true is a *data point* that must be accommodated by any viable theory of free will. That is, to answer "no" to the correlation question is to commit oneself to a desideratum on any viable theory of free will, namely: it must include at least one necessary condition which cannot be satisfied by a normal human living in a world at which determinism is true. By contrast, those who answer "yes" thereby reject that there is any such desideratum on a viable account of free will. In other words, determinism scenarios reveal that people not only have different intuitions about what free will is, but they also disagree about the standards by which a theory of free will is to be judged.

Despite the dialectical significance of the "yes" and "no" answers to the correlation problem, there is no standard jargon which unequivocally tracks the modal views associated with them. To remedy this, I have introduced new terminology (Mickelson 2012, 2015a, 2015b, 2017, 2019a, 2019b, 2021a). To give a "no" response to the correlation problem is to claim that human free will and determinism-related factors are *incompossible*, that is, it is metaphysically impossible for them to co-exist. In light of this, I say that those who answer "no" to the correlation problem are proponents of an *incompossibility solution* to the correlation problem, while those who answer "yes" endorse a *compossibility solution*. Correspondingly, *incompossibilism* is the view that the incompossibility solution to the correlation problem is correct, and *compossibilism* is the view that the compossibility solution is correct.

2 The Explanation Problem

Incompossibilism answers one problem only to raise another: unlike the compossibilist, the incompossibilist must explain or account for the incompossibility of human freedom and determinism-related factors. To solve this *explanation problem*, the incompossibilist must answer the following questions:

> **The Unmet Condition Challenge** (**E1**): Which necessary condition C on free will cannot be satisfied by a normal human when determinism is true?

The Condition Underminer Challenge (**E2**): What prevents a normal human from satisfying condition C when determinism is true?[7]

Although the E1/E2 distinction is present in contemporary discussions of free will, free-will specialists have not made an effort to track this distinction or that between an incompossibility solution to the correlation problem and extant solutions to E2 of the explanation problem. Presumably, this is at least partly because their understanding of the problem of determinism has been shaped and constrained by the distinctive jargon and narratives of the classical analytic paradigm. Given these historical considerations, let us begin by looking at the bipartite explanation problem through the narrow lens of the classical analytic paradigm.

2.1 Classical Analytic Solutions to the Correlation and Explanation Problems

Within the classical analytic paradigm, the questions associated with E1 and E2 were not distinguished from a "no" answer to the correlation problem, and it's easy to understand why. Given the background assumptions of that paradigm, a complete solution to the explanation problem follows directly from an incompossibility solution to the correlation problem. How so? Classical analytic theorists agreed that free will is (or requires) an *ability to do otherwise*. Since it is standard practice to attach the term 'classical' to view which presuppose that free will is an ability to do otherwise, it is natural to refer to the incompossibilist who accepts the classical view of free will as a *classical incompossibilist*. The classical incompossibilist must interpret her solution to the correlation problem as providing key information about what this freedom-relevant ability to do otherwise amounts to. Specifically, the classical incompossibilist commits to the view that free will, that is, the ability to do otherwise, cannot be exercised in which one unique future is made inevitable by determinism-related factors (given the facts of the past). Free will is a more robust ability to do otherwise, roughly an ability to "settle" which of the multiple futures available to her at the moment of action, in the forking-paths timeline of her life, comes to pass; in counterfactual terms, an actor exercises the relevant ability to do otherwise in performing A only when it is true of the actor that, holding fixed the laws and the facts of the past prior to the action, the actor still *could have done otherwise* than perform A, that is, she could have *avoided* or *refrained* from doing what she actually did and done something else instead (e.g. van Inwagen 1983, 2017; Balaguer Ch. 9 this volume).

Notably, classical analytic theorists also took for granted the truth of *anthropocentric possibilism*, the view that it is at least metaphysically possible for an ordinary human to exercise free will (e.g. van Inwagen 1983; Clarke 2003; Vihvelin 2013).[8] Any account of free will that is consistent with both incompossibilism and anthropocentric possibilism is now commonly classified as a *broadly libertarian account* of free will.[9] This means that all incompossibilists working in the classical analytic paradigm were committed to a broadly libertarian solution to E1.[10] I will hereafter refer to the general solution to E1 endorsed by classical incompossibilists, according to which some libertarian interpretation of the ability to do otherwise is correct, the *classical libertarian solution* to E1.[11]

According to the classical libertarian solution to E1, it is metaphysically possible for a normal human to exercise free will only if there is some type of *actual-sequence leeway* in the evolution of the universe. There is actual-sequence leeway in the evolution of the universe only if there is at least one point in time at which, holding fixed the laws and facts of the past, more than one future might unfold. That is, more than one future is *accessible* from some point in the actual timeline, just as we see in van Inwagen's "forking paths" depiction of

indeterminism (in van Inwagen's diagram seen earlier). Recall that determinism, by definition, is true only when there is literally zero actual-sequence leeway in the evolution of the physical world, that is, when exactly zero "forking paths" are permitted in the actual timeline. This means that if determinism were true, certain determinism-related factors would rule out the sort of actual sequence leeway – *whatever type that may be* – that an exercise of free will requires (given the libertarian take on the notion of an ability to do otherwise). Within the classical analytic paradigm, then, an incompossibility solution to the correlation problem implies the classical libertarian solution to E1, which in turn implies the *classical incompatibilist solution* to E2: when determinism is true, it is *owing to determinism-related factors* that people cannot exercise the ability to do otherwise that an act of free will requires. For ease of reference, I will hereafter refer to this line of reasoning – from incompossibilism, via the classical libertarian solution to E1, to the classical incompatibilist solution to E2 – as the *classical bridge inference*.[12]

Now that we have a better grasp on the classical analytic characterization of the problem of determinism, we can understand why classical analytic theorists believed – as their Compatibility Problem narrative tells us – that the problem admits only two (mutually exclusive and collectively exhaustive) candidate solutions. On the one hand, classical analytic theorists could adopt an incompossibility solution to the correlation problem which, given the background assumptions of the classical analytic paradigm, came theoretically loaded – via the classical bridge inference – with the classical libertarian solution to E1 and the classical incompatibilist solution to E2. This complex position is now known as *classical incompatibilism*. On the other hand, the classical analytic theorists could embrace a compossibility solution to the correlation problem and proclaim that the explanation problem is a pseudo-problem generated by a false (incompossibility) solution to the correlation problem and, so, all candidate solutions to the explanation problem are wrong. The latter position is now known as *classical compatibilism*. For those working within the classical analytic paradigm, then, it was quite sensible to summarize the problem of determinism as a debate about whether classical compatibilism or classical incompatibilism is true.

2.2 Contemporary Solutions to E1 of the Explanation Problem

The classical analytic paradigm has fallen out of favor, largely because many philosophers are unwilling to grant its presumptions that free will is/requires an ability to do otherwise and that anthropocentric possibilism is true.[13] Indeed, a growing number of contemporary philosophers explicitly reject both anthropocentric possibilism and the classical definition of 'free will'. For example, Galen Strawson (1986, 1994, 2002, 2008, 2011) and Derk Pereboom (2001, 2014) are incompossibilists who reject the classical libertarian solution to E1, favoring instead a type of *source solution* to E1 according to which free will requires one to be the genuine *source* of one's action. Although their accounts of freedom-relevant sourcehood are substantively different, Pereboom and G. Strawson agree that it is metaphysically impossible for a normal human to satisfy the necessary source condition on free will.[14] We might say, then, that G. Strawson and Pereboom are each *source incompossibilists*, insofar as each embraces a strict source solution to E1 and, moreover, each is an *anthropocentric source impossibilist*, insofar as each embraces the view that it is metaphysically impossible for an ordinary human to satisfy the source condition on free will. However, whereas G. Strawson's understanding of the source condition leads him to endorse unqualified *impossibilism*, that is, the unqualified view that it is metaphysically impossible for any being to exercise free will, Pereboom's understanding of the source condition leads him to reject impossibilism (e.g. Pereboom 2001, p. 132).

In addition, some philosophers have suggested a type of hybrid solution to E1, proposing that an ability to do otherwise is required to satisfy the source condition on free will. For example, Robert Kane's "self-forming actions" account of free will seems to be a hybrid solution of this kind (e.g. Kane 1996, 2004, 2016, 2019; see also Balaguer CH. 9 this volume). As such, it would now be question-begging to frame the free-will debate, as classical analytic theorists once did, against the background assumptions that anthropocentric possibilism is true and that some version of the classical (ability-to-do-otherwise) characterization of free will is correct.

The fact that contemporary incompossibilists disagree about the nature of free will in a more fundamental way than their classical predecessors makes it easier to understand why contemporary free-will theorists struggle to provide a precise yet reasonably uncontroversial characterization of *free will* (and/or definition of 'free will') for use in their debate. The more precise one's proposed characterization, the more likely that one camp or another in the debate will find it question-begging; the more generic the proposal, the more likely that the referent of 'free will' will not be fixed securely enough to ensure that the interlocutors in the debate are disagreeing about the same thing (rather than having a mere verbal dispute in which two groups talk past each other because they are using the same phrase 'free will' to pick out different things).

There is no simple fix for this dialectical difficulty, but neo-classical and non-classical theorists have generally agreed to fix the referent of 'free will' by saying that free will is a type of *control* or *up-to-one-ness* that is necessary for moral responsibility, such that a person is morally responsible for an action A only if she exercises this type of control or exhibits this type of up-to-one-ness when performing A. Since it would be easy to find objections to any attempt to give a dialectically neutral characterization of free will (or definition of 'free will'), I will not attempt to provide one here; I leave readers to weigh the pros and cons of popular working definitions of 'free will' for themselves.

2.3 Contemporary Solutions to E2 of the Explanation Problem

Philosophers who are unaware of recent shifts away from the classical analytic paradigm may find it strange, even absurd, to suggest that there is room for a substantive and philosophically interesting debate about what *explains* and/or *accounts for* the lack of free agents in worlds at which determinism is true. Readers in this camp might wonder: Assuming that incompossibilism is true, isn't it just obvious – so obvious that it may go unstated – that certain *determinism-related factors* play at least some role in making people unfree when determinism is true? Does it really matter in the end whether the incompossibilists disagree about the fine-grained details regarding *which* determinism-related factors – the causation, the laws of nature, a conjunction of causal factors and facts about the past, etc. – pose a threat to human freedom, or whether determinism-related factors preclude free will *on their own* or only in conjunction with some other factors beyond our control?

Readers drawn to such questions must be reminded that the contemporary free will debate is no longer constrained by the assumptions and methods of the classical analytic paradigm. Once one abandons the classical analytic stipulation that 'free will' denotes an ability to do otherwise, the classical bridge inference is no longer a viable way to close the "explanatory gap" between an incompossibility solution to the correlation problem and the classical incompatibilist solution to E2 of the explanation problem. Indeed, for all that is stated by incompossibilism, it might simply be a *brute fact* that no one acts freely when determinism is true.[15] Of course, few would accept that the incompossibility of free will and determinism-related

factors is a brute fact; assuming incompossibilism is true, there must be some *better explanation* for its truth than that. But what *is* that better explanation? This brings us to the central point of this section: influential incompossibilists can and do disagree about what makes people unfree when determinism is true. That is, contemporary incompossibilists, unlike their classical predecessors, substantively disagree about the correct solution to E2.

The suggestion that the literature already includes a variety of philosophically interesting and conflicting solutions to E2 may take some readers by surprise. After all, philosophers generally speak as though all incompossibilists agree that determinism-related factors keep people from acting freely when determinism is true, but are divided about whether such factors preclude free will *because they keep people from exercising the ability to do otherwise* or, rather, such factors preclude free will *because they keep people from being the source of their own actions* (e.g. Fischer and Ravizza 1998, p. 151, Kane 2002, p. 6; Griffith 2017, p. 2; Vihvelin 2022). Unfortunately, the existence and contours of this lively in-house dispute among incompossibilists have been obscured in part because the dispute cannot be adequately characterized using the jargon and narratives of the classical analytic paradigm.

For purposes of categorizing extant solutions to E2, it is helpful to distinguish between two basic categories of factors which are, arguably, *beyond one's control*. First, there are two types of actor-extrinsic factors. On the one hand, there are states of physical world outside (i.e. external to the properties which constitute) a given actor. These actor-extrinsic factors include both aspects of one's immediate environment (e.g. the physical properties of the room one is sitting in) and complete states of the world in the remote past, for example, states which were obtained before the first human was born (e.g. there were no humans when dinosaurs roamed the Earth). On the other hand, there are actor-extrinsic factors which govern or otherwise account for the way in which the universe evolves from one state to another over time. For example, *determinism-related factors* are those factors which, given the state of the physical universe at one time, ensure the fine-grained state of the physical universe at the next moment--and every moment thereafter, until the end of time. The specific nature of the factors which account for the evolution of the universe is a matter of debate, for example, they may be causal relations, laws of nature, and so on. Second, there are features of an actor which may be (arguably, at least) beyond an actor's control. These include the properties which an actor has at a given moment (e.g. one's precise brain state at a given moment), features an actor has during their entire lifespan (e.g. one's genetic endowment), as well as an actor's essential properties, if there are such things (e.g. those properties which establish an actor's cross-world identity, such that we can posit that an actor in one possible world is – or at least has a counterpart in – some other possible world).

We might introduce a completely new set of labels to track the three mutually exclusive and collectively exhaustive classes of factors described earlier, but I will instead import and refine a few terms that have become commonplace in discussions of the paradox of moral luck (e.g. Hartman Ch. 10 this volume). This choice reflects my conviction that the problem of determinism and the paradox of moral luck are best understood as two rhetorically distinct frameworks for investigating *the problem of free will*, roughly the problem of identifying what free will is and whether we (i.e. ordinary humans) have it.[16] Like the problem of determinism, the paradox of moral luck is grounded partly in the observation that our normal practices of blame and punishment seemingly presuppose that we have the type of control required for basic-desert moral responsibility (McCormick Ch. 24 this volume), and yet close examination of individual actions indicates that no human has that type of control – rather, every human action is settled mostly – if not entirely – by factors beyond our control.[17] Moral-luck theorists have normalized the practice of using 'luck' as shorthand for the more cumbersome phrase

"factors beyond one's control" (Hartman 2017, pp. 23–31; Anderson 2019; Statman 2019), and I will follow that convention here. Inspired by Nagel (1976, 1979, 1986), I will speak of three basic types of luck: causal luck, circumstantial luck, and constitutive luck.[18]

For present purposes, let us say that a person is subject to *causal luck* when one lacks control over the factors which settle how the world evolves from one state of affairs to another (e.g. causal relations, causal laws, laws of nature, or the like). Let us say that one is subject to *circumstantial luck* when one lacks control over the non-causal/agent-extrinsic states of affairs (e.g. states of the universe prior to one's birth). Finally, one is subject to *constitutive luck* when one lacks control over one's own constitutive properties (e.g. one's genetic endowments with regards to intelligence and personality, or that one is an ordinary human rather than some other type of subhuman or superhuman creature).[19] Using these three categories of luck, we can easily identify a variety of extant solutions to E2: *causal luck solutions*, *circumstantial luck solutions*, *constitutive luck solutions*, and *hybrid solutions*. Let's consider each in turn.

A strict *causal luck solution* to E2 proposes that it is impossible for people to act freely when determinism is true because there are certain nomological and/or causation-related factors which make it the case that no human can exercise free will. The classical incompatibilist ("owing to determinism") solution to E2 may be classified as a type of causal luck solution. That said, there is plenty of room to wonder what this classical causal luck solution really amounts to. Philosophers commonly speak as though determinism per se precludes free will. Taken literally, however, this is an untenable response to E2. Why? Determinism has the ontological status of a proposition, and this proposition merely provides a *description* of a certain kind of world. There is little reason to suppose that a description (even if true) could itself pose a threat to free will; it is not the description, but *what it describes*, that has the ontological standing to preclude the exercise of free will (e.g. Hermes and Campbell 2012; Hermes 2013).

Philosophers who contend that certain evolution-governing factors *described by* determinism have the ontological standing required to undercut human freedom (as opposed to the thesis of determinism itself) carry the dialectical burden of specifying *which* evolutionary factors do the freedom-undermining work when determinism is true. Philosophers often pinpoint the *deterministic* quality of causation (laws of nature, or the like) which obtain when determinism is true as a specific threat to free will (e.g. Pereboom 2001, 2014; Sartorio 2016). Notably, though, the laws of nature may be deterministic in a perfectly standard sense of the term 'deterministic' even when the type of zero-leeway determinism described by James and depicted by van Inwagen (above) is false (e.g. Stone 1998; Dennett 2003; Sehon 2010; Mickelson 2012, 2019a, 2019b).[20] Other potentially freedom-relevant features of the laws of nature described by determinism are their *strength*, that is, that they are "strong" rather than "weak" (Perry 2004), and their *potency*, that is, that such laws are "unconditional" rather than "conditional" (Mickelson 2019a).[21] Given these distinctions, there is ample dialectical space for a lively in-house debate among the subset of incompossibilists who endorse a causal luck solution to E2 – a group we might call *causal luck incompossibilists* – to disagree about which determinism-related factors (if any) are antagonistic to free will. Given that some determinism-related factors are present even when determinism is false, such disagreements may have very interesting implications for the general project of assessing which account of free will is best (Mickelson 2019a and "Hard Times for Hard Incompatibilism," ms.).

Incompossibilists who reject broadly libertarian solutions to E1 may – and arguably must[22] – categorically reject causal luck solutions to E2. For example, G. Strawson is an impossibilist, and *a fortiori* an incompossibilist, who rejects the classical solutions to E1 and E2. With his Basic Argument, G. Strawson argues that we ("the folk") intuitively accept that there is an "ultimate starting point" source condition on free will, but the type of "buck stops here"

"factors beyond one's control" (Hartman 2017, pp. 23–31; Anderson 2019; Statman 2019), and I will follow that convention here. Inspired by Nagel (1976, 1979, 1986), I will speak of three basic types of luck: causal luck, circumstantial luck, and constitutive luck.[18]

For present purposes, let us say that a person is subject to *causal luck* when one lacks control over the factors which settle how the world evolves from one state of affairs to another (e.g. causal relations, causal laws, laws of nature, or the like). Let us say that one is subject to *circumstantial luck* when one lacks control over the non-causal/agent-extrinsic states of affairs (e.g. states of the universe prior to one's birth). Finally, one is subject to *constitutive luck* when one lacks control over one's own constitutive properties (e.g. one's genetic endowments with regards to intelligence and personality, or that one is an ordinary human rather than some other type of subhuman or superhuman creature).[19] Using these three categories of luck, we can easily identify a variety of extant solutions to E2: *causal luck solutions*, *circumstantial luck solutions*, *constitutive luck solutions*, and *hybrid solutions*. Let's consider each in turn.

A strict *causal luck solution* to E2 proposes that it is impossible for people to act freely when determinism is true because there are certain nomological and/or causation-related factors which make it the case that no human can exercise free will. The classical incompatibilist ("owing to determinism") solution to E2 may be classified as a type of causal luck solution. That said, there is plenty of room to wonder what this classical causal luck solution really amounts to. Philosophers commonly speak as though determinism per se precludes free will. Taken literally, however, this is an untenable response to E2. Why? Determinism has the ontological status of a proposition, and this proposition merely provides a *description* of a certain kind of world. There is little reason to suppose that a description (even if true) could itself pose a threat to free will; it is not the description, but *what it describes*, that has the ontological standing to preclude the exercise of free will (e.g. Hermes and Campbell 2012; Hermes 2013).

Philosophers who contend that certain evolution-governing factors *described by* determinism have the ontological standing required to undercut human freedom (as opposed to the thesis of determinism itself) carry the dialectical burden of specifying *which* evolutionary factors do the freedom-undermining work when determinism is true. Philosophers often pinpoint the *deterministic* quality of causation (laws of nature, or the like) which obtain when determinism is true as a specific threat to free will (e.g. Pereboom 2001, 2014; Sartorio 2016). Notably, though, the laws of nature may be deterministic in a perfectly standard sense of the term 'deterministic' even when the type of zero-leeway determinism described by James and depicted by van Inwagen (above) is false (e.g. Stone 1998; Dennett 2003; Sehon 2010; Mickelson 2012, 2019a, 2019b).[20] Other potentially freedom-relevant features of the laws of nature described by determinism are their *strength*, that is, that they are "strong" rather than "weak" (Perry 2004), and their *potency*, that is, that such laws are "unconditional" rather than "conditional" (Mickelson 2019a).[21] Given these distinctions, there is ample dialectical space for a lively in-house debate among the subset of incompossibilists who endorse a causal luck solution to E2 – a group we might call *causal luck incompossibilists* – to disagree about which determinism-related factors (if any) are antagonistic to free will. Given that some determinism-related factors are present even when determinism is false, such disagreements may have very interesting implications for the general project of assessing which account of free will is best (Mickelson 2019a and "Hard Times for Hard Incompatibilism," ms.).

Incompossibilists who reject broadly libertarian solutions to E1 may – and arguably must[22] – categorically reject causal luck solutions to E2. For example, G. Strawson is an impossibilist, and *a fortiori* an incompossibilist, who rejects the classical solutions to E1 and E2. With his Basic Argument, G. Strawson argues that we ("the folk") intuitively accept that there is an "ultimate starting point" source condition on free will, but the type of "buck stops here"

sourcehood we intuitively want – and generally take ourselves to have – could be achieved only be a *causa sui*.[23] Satisfying the sourcehood condition, in other words, would require an act of ex nihilo self-creation. However, the notion of self-creating ex nihilo implies the existence of a *nothing-self*, which is a contradiction in terms. Clearly, the notion of ex nihilo self-creation is incoherent *irrespective of what the world is like* (e.g. irrespective of what type of laws of nature obtain or what the remote past was like). In other words, the Basic Argument tells us that all causal luck and circumstantial luck solutions to E2 are false; it is specifically *constitutive luck*, and not any and every sort of luck around, that "swallows everything" (Mickelson 2019b). Among other things, this means that philosophers who are still working on the outdated classical analytic assumption that a mere incompossibility solution to the correlation problem guarantees that a causal luck ("incompatibilist") solution to E2 – or, equivalently, that any argument for incompossibilism is an argument for incompatibilism (a.k.a. causal luck incompossibilism) – is guilty of begging the question against someone (we might call them a *constitutive luck impossibilist*) such as G. Strawson.

There seems to be little sympathy in the literature for a *strict circumstantial luck solution* to E2, that is, a solution which says that it is impossible to satisfy the condition on free will named in E1 when determinism is true strictly because people lack control over certain (non-causal, agent-extrinsic) states of affairs which obtained prior to their actions.[24] However, philosophers often combine circumstantial luck with another type of luck to create a hybrid solution to E2. For example, some philosophers promote a *causal/circumstantial luck hybrid solution* according to which it is the combination of our lack of control over the laws and circumstances in the past which makes us unfree. For example, Pereboom appears to be in this camp, as he contends that normal humans would lack free will when determinism is true because in such conditions one's actions are deterministically caused *by factors beyond their control*, where it appears that the additional "factors" he is alluding to are certain past states of the world over which the actors exercised no control (Pereboom 2001, 2014). Constitutive luck solutions to E2 may also come in hybrid form, and some philosophers have suggested that circumstantial/constitutive luck hybrid solutions are more compelling than any type of causal luck solution (e.g. Nagel 1986, pp. 113–114; Latus 2001; Levy 2011, p. 96).[25]

Summing up, we have seen that classical analytic theorists conceived of the free-will debate as an in-house dispute among anthropocentric possibilists over the correct understanding of *the ability to do otherwise*, with the classical compossibilists adopting one view and classical incompossibilists adopting another. The classical libertarian solution to E1 fit perfectly with the classical incompatibilist solution to E2, which accounts for the fact that all classical incompossibilists were also classical incompatibilists. Since all classical compossibilists rejected both incompossibilism and the classical incompatibilist solution to E2, it made sense for classical analytic theorists to propose that the fundamental divide in the free will debate is that between (classical compossibilist) compatibilists and (classical broadly libertarian) incompatibilists. However, the free will debate has outgrown the bipartite compatibilism/incompatibilism taxonomy that we inherited from the classical analytic theorists. In the contemporary debate, philosophers are invited to select from a much wider range of fundamentally different positions.

2.4 Testing the Quality of Candidate Solutions to the Explanation Problem

The observation that incompossibilists can and do disagree about the correct solution to explanation problem sheds new light on the overall dialectic of the free-will debate. Teasing apart E1 and E2 makes it easier to see that an incompossibilist who aims to solve the

explanation problem must provide a solution to E2 that fits with his proposed solution to E1. By "fit," I mean that whatever is singled out as the "threat" to free will in one's solution to E2 must be the kind of thing that could keep a person from satisfying the necessary condition on free will named in one's solution to E1. For example, we have seen that the internal coherence of the classical incompatibilist solution to E2 of the explanation problem was assured, via the classical bridge inference, by the classical libertarian solution to E1. Likewise, G. Strawson's suggestion that people have an incoherent "ultimate starting point" conception of source-hood fits well with the view that constitutive luck (i.e. the lack of relevant control over one's constitutive properties) and not contingent determinism-related factors is what makes it impossible for a person to satisfy the source condition on free will when determinism is true.[26]

Stepping back, the fact that incompossibilists disagree about what free will is and what undermines it implies that incompossibilists also disagree about what incompossibilist intuitions are tracking. We must resist the temptation (encouraged by the lingering influence of classical analytic narratives) to assume that incompossibilist intuitions are rational insofar as they track the freedom-undermining effects of determinism-related factors, for some incompossibilists contend that they track the threat posed by the toxic combination of determinism-related factors and circumstantial factors, others contend that these intuitions track the menacing nature of constitutive luck alone, and the list goes on. These rival accounts of what makes incompossibilist intuitions rational and incompossibilism true cannot all be right, but they may all be wrong – after all, incompossibilism may be false and all incompossibilist intuitions may be faulty.

It should now be clear that embracing incompossibilism is just a starting point for a person with incompossibilist intuitions. His project is not done until he tells his rivals – compossibilists, agnostics/neutral-inquirers, and rival incompossibilists alike – a plausible and detailed story about what free will is (i.e. by answering E1) and precisely what undermines it when determinism is true (i.e. by answering E2). Indeed, there seems no way to assess – therefore no reason to *accept* – that incompossibilism is true and/or that incompossibilists intuitions (as opposed to compossibilist intuitions) are truth-tracking until incompossibilism is coupled with a specific solution to the explanation problem. Put another way, incompossibilism is a useful rhetorical landmark in a complicated dialectic, but it is not among the viable endgame positions in the free-will debate.[27]

3 The "Incompatibility" Relations of Incompatibilism

Since the early stages of the classical analytic period in the free-will debate, free-will theorists have been conflating relations which are as fundamentally different in kind as correlation and causation under the term 'incompatibility.' Just as there is no single incompatibility relation picked out by the term 'incompatibility,' the term 'compatibility' has become an umbrella term for a variety of non-incompatibility relations. While the failure to track these distinctions was arguably little more than a bit of bad housekeeping for those working in the classical analytic paradigm, these distinctions are indispensable to those working with a more contemporary characterization of the debate. In this section, I review the relations currently conflated under the label 'incompatibility,' and thereby expose the substantive ambiguity of the terms 'incompatibilism' and 'compatibilism.'

3.1 Relevance Relations versus Correlation Relations

Outside of academic philosophy, people speak of an "incompatibility" relation between things when they wish to indicate that those things cannot co-exist in harmony in virtue of a *conflict*

rooted in their respective natures. So understood, incompatibility is a type of antagonistic *relevance relation*. To say that a relevance relation holds between two relata is to say that the relata are connected in such a way that the existence and/or specific properties of one relatum (e.g. its causal properties, its truth-value) are *relevant to* the existence of and/or specific properties of the other. As I am characterizing it, the class of relevance relations is very broad and includes causal relations (assuming a strong/non-Humean account of causation), settling relations, fixing relations, non-causal determining relations, grounding relations, backing relations, and explanatory "because" relationships.[28]

Relevance relations may be fruitfully contrasted with another extremely broad class of association relations, namely *correlation relations*. As suggested by the etymology of the term 'correlation' – it is derived from a combination of the Latin *cor-* ("together") and *relatio* ("relation") – a correlation is a type of mutual relationship or connection between two or more things which may or may not be underwritten by a relevance relation. That is, unlike a relevance relation, the type of co-occurrence or co-variation associated with correlation – no matter how regular or useful for making predictions – may be a complete fluke. While the strength of a correlation is often described in statistical terms, the generic class of correlation relations is quite diverse. For example, the non-explanatory covariance relation of *supervenience* is a mere correlation relation which is sometimes contrasted with the explanatory relevance relation of *superdupervenience* (e.g. McPherson 2021). When two things stand in a relevance relation, there will also be a correlation relation between them; however, a correlation relation may or may not be underwritten by a relevance relation. Because of this asymmetry, it is always a mistake to infer that a relevance relation holds between two things on the meager grounds that they are correlated. As such, the distinction between correlation relations and relevance relations allows us to understand the *cum hoc, ergo propter hoc* fallacy in its broadest terms. We will return to the topic of correlation relations in a moment.

3.2 Three Relevance Relations: Conceptual, Metaphysical, and Logical Incompatibility

In standard discourse, *incompatibility* is an antagonistic relevance relation, that is, it is not a perhaps spurious correlation relation. Familiar examples of incompatibility, in the standard relevance-relation sense, include: *incompatible medications* (which chemically interact in a way that reduces the wanted medicinal effects or brings about negative effects for the patient), *incompatible roommates* (who have conflicting personalities), and *genetic incompatibility* (which indicates that the genetic properties of two individuals make it impossible for them to produce viable – live, non-sterile, etc. – offspring). Based on the way free-will specialists speak – that determinism *threatens* (Baker 2006, p. 313), *conflicts with* (Kane 1999, p. 218; Chisholm 1964/2009, p. 24; Ayer 1954/2009, p. 15), *rules out* (McKenna and Pereboom 2016, p. 169), *undermines* (Nahmias et al. 2006, p. 40), is *at odds with* (Pereboom 2005, p. 240), *precludes* (Mele 2006, p. 189), *vitiates* (Fischer 1994, p. 159), is *menacing* to (Haji and Cuypers 2006), is *a problem for* (Campbell 2011, p. 23), and *cannot be reconciled with* (Dennett and Caruso 2021, p. 5) free will and/or moral responsibility[29] – the non-specialist might reasonably suppose that incompatibilism is named after some variation of the standard antagonistic, relevance-relation notion of incompatibility. On this interpretation, 'incompatibilism' names the view that determinism-related factors are incompatible with – in the sense that they destroy, rule out, preclude, undermine, make impossible, conflict with – free will, that is, it names a broadly causal luck solution to E2 of the explanation problem. Notably, most philosophers consider this type of relevance-relation incompatibility to be at the heart of the "traditional" incompatibilist position (e.g. Pereboom 2001, pp. 128–129; Vihvelin 2008, 2013; McKenna 2010, p. 432; Levy 2011, p. 1, n.1; McKenna and Pereboom 2016, p. 151).

Incompatibility relations come in a variety of species, some of which may be distinguished by the ontological status of the relata. For example, when philosophers talk about the *conceptual incompatibility* of free will and determinism, they are talking about an incompatibility relation that holds between *concepts* or the *conceptual content* of certain statements (e.g. van Inwagen 1983, pp. 1, 66, 87; Nahmias et al. 2006, p. 30; Mele 2012, p. 434). In these discussions, context clearly indicates that the standard relevance-relation notion of incompatibility is at play. For example, the concepts BACHELOR and MARRIED are conceptually incompatible in the relevant sense. Unpacking the content of BACHELOR, we find that a necessary condition on being a bachelor is *being unmarried*, and this reveals a direct tension in the content of the concepts BACHELOR and MARRIED – a tension which accounts for (explains, grounds, or the like) the fact that it is impossible for anything to *be* a married bachelor. Likewise, to say that the concept of FREE WILL is incompatible with DETERMINISM is to say that there is a conflict rooted in the content of these two concepts from which it may be derived that it is impossible to exercise free will when determinism is true. It is unsurprising that philosophers working in the wake of the classical analytic paradigm would focus on the issue of conceptual incompatibility, for – as the discussion of the classical bridge inference in the previous section makes plain – the classical incompossibilists' proposal that some broadly libertarian interpretation of the ability to do otherwise is a necessary condition on free will commits them to the conceptual incompatibility of FREE WILL and DETERMINISM.

Although conceptual incompatibility is an interesting relation, it is arguably not a basic one. We can unpack this relation in either logical or metaphysical terms. In logical terms, we might flesh out the conceptual incompatibility of BACHELOR and MARRIED by saying that the proposition *that S is a bachelor* and the proposition *that S is married* are *logically incompatible*: the truth of the former proposition (in virtue of its meaning) would preclude the truth of the latter (in virtue of its meaning), and vice versa. Put another way, the statement "S is a married bachelor" is self-contradictory given the meaning of the terms in the statement. For ease of reference, I will call a logical incompatibility relation that holds necessarily a *strict logical incompatibility relation* (since the term 'strict' is often used in logical contexts to connote necessity).[30]

Alternatively, we might cash out conceptual incompatibility as a *metaphysical incompatibility* relation. Continuing with our example, a thing cannot "fall under" the concepts BACHELOR and MARRIED simultaneously: any person who satisfies the necessary conditions on being a bachelor is someone who has the property of being unmarried, and it is impossible for someone who has the property of being unmarried to simultaneously have the property of being married. By contrast, BACHELOR and HAPPY are not conceptually incompatible, which we can unpack by saying that it is metaphysically possible for a single thing to exist which simultaneously "falls under" or "answers to" both concepts.

3.3 Two Correlation Relations: Inconsistency and Incompossibility

The three incompatibility relations discussed so far belong to the class of non-spurious *relevance relations*, a class I have contrasted with perhaps spurious *correlation relations*.[31] With this distinction drawn, let us further flesh out the distinction by considering different kinds of correlation relations which are routinely picked out by the terms 'incompatible' and 'incompatibility' in the contemporary free-will literature: *metaphysical incompossibility* and *strict logical inconsistency*.

Metaphysical incompossibility is a type of correlation relation and, as with any other correlation relation, it may be *spurious*. Two subtypes of spurious correlation are worth noting,

what I will call *indirect correlation* and *trivial correlation*. A correlation relation is *indirectly spurious* when the correlation between phenomena is brought about or otherwise ensured, via independent routes, by some third variable (a.k.a. a confounding variable). A stock example of an indirectly spurious correlation relation is the strong correlation between ice cream sales and murder rates: as ice cream sales rise and fall in a given region, so do the murder rates there. However, this is not because eating ice cream *causes* people to commit murder or that murdering inclines people toward eating ice cream. Rather, the rise in ice cream sales and murder rates is only *indirectly related* by a third variable which brings about each (via independent causal mechanisms), namely: the rise in outdoor temperatures. On the other hand, trivially spurious correlations – such as that between the number of people who drown by falling into pools each year during a given period and the number of films that Nicolas Cage appeared in each year during that period – are mere coincidence; there is no hidden third variable or other underlying story which accounts for and/or explains them.

Spurious metaphysical incompossibility also comes in both indirect and trivial forms. A salient example of *indirect incompossibility* is evident in the following view about the relationship between God's foreknowledge and human free will. Assume for a moment that determinism is true and that free will is metaphysically incompatible with – directly destroyed by, precluded by, undermined by – the type of causal laws described by determinism (i.e. grant that some type of causal luck solution to E2 is correct). Assume also that God exists. From these assumptions, one could plausibly argue that God knows (given his knowledge of the laws of nature and facts about the early state of the universe in which we live) what will unfold in the future timeline of this universe, including every action a given person performs. However, it does not follow from this that God's foreknowledge is a freedom-undermining factor of the world. One can reasonably hold that God's foreknowledge is *incompossible* with free will, but only spuriously so. For example, one may identify some third variable, such as the causal relations described by determinism, as the feature of the world which directly destroys free will and (independently) accounts for God's foreknowledge. Here, God's foreknowledge – even if we grant that it is *a factor beyond our control* – is not metaphysically/explanatorily relevant vis-à-vis the fact that no one acts freely when determinism is true, just as the rise in murder rates is not causally/explanatorily relevant to the rise of murder rates in the example earlier.[32] In short, one may plausibly affirm that God's foreknowledge is (spuriously) incompossible with human free will while denying that these phenomena are metaphysically incompatible.

Metaphysical incompossibility may also be completely trivial. A spurious incompossibility relation is *trivial* when it holds between two phenomena which bear no direct nor indirect relevance relationship to one another. For example, fluffy kittens and round squares are incompossible, but only trivially so: the existence of fluffy kittens does not keep round squares from existing (or vice versa). In this case, the incompossibility follows trivially from the metaphysical impossibility of round squares alone. Although free-will experts have made no consistent effort to track the distinction between metaphysical incompatibility and spurious incompossibility in recent years, the value of the distinction should be plain enough to open-minded inquirers. For example, only someone with a good grasp of this distinction is in position to grasp the upshot of G. Strawson's Basic Argument. As noted in the previous section, the Basic Argument aims to show that free will is metaphysically impossible because people's actions are settled by *constitutive* properties of the agent that are beyond that agent's control, that is, free will is impossible due to constitutive luck. It follows *a fortiori* from the impossibilist conclusion of the Basic Argument that free will is trivially incompossible with all determinism-related factors (e.g. deterministic laws of nature) and, in the same uninteresting way, that

free will is trivially incompossible with fluffy kittens, dirty diapers, and lilac bushes in bloom. So, if the Basic Argument is sound, then determinism-related factors – despite being among the factors beyond human control – do not even partly account for or explain the lack of free agents in worlds at which determinism is true; determinism-related factors are completely irrelevant to – and, hence, are not metaphysically incompatible with – free will. In sum, the Basic Argument promotes a solution to E2 which implies that causal luck solutions (including the classical incompatibilist solution) to E2 are categorically false – for this argument concludes that determinism-related factors are *trivially spuriously* incompossible with free will, which is something that all causal luck solutions to E2 deny (since the latter, by definition, propose that a *non-spurious* antagonistic relevance relation holds between these relata).

Corresponding to the sharp distinction between the metaphysical relations of metaphysical incompossibility and metaphysical incompatibility, we may draw a distinction between the logical relevance relation of *strict logical incompatibility* and its correlation counterpart *strict logical inconsistency* (a.k.a. *logical non-compossibility, logical incompossibility*).[33] Here, I assume a common non-relevance definition of 'logical inconsistency': Statements are *logically inconsistent* if and only if they cannot have the truth-value *true* at the same time.[34] In the language of first-order propositional logic, we may say that propositions *p* and *q* are logically inconsistent if and only if ~(p · q) is true, or, equivalently, that *p* materially implies the negation of q, that is, that (p → ~ q).[35] The latter way of expressing the logical inconsistency is relation is particularly useful for our purposes since it is uncontroversial and widely known that material implication is *not* a relevance relation (e.g. Mares 2020). In short, to say that *p* and *q* are logically inconsistent is to give us information about how the truth-values of *p* and *q* covary; it tells us nothing about *why* the truth-values covary in this way. For example, such claims do not tell us that the inconsistency is due to some kind of *conflict* between propositions *p* and *q*, for example, that their inconsistency is due to an underlying (syntactic and/or semantic) *incompatibility*.

Like metaphysical incompossibility, a logical inconsistency relation may or may not be spurious. For example, given the spurious correlation between ice cream sales and homicide rates (noted earlier), the material conditional "If the rate of ice cream sales rises during period P, then the rate of homicides increases during period P" is true, but it would be a mistake to think that this material conditional claims that the rise of ice cream sales is directly relevant to (accounts for, causes, etc.) the rise in the number of homicides during P. Moreover, an inconsistency relation between two statements (and/or propositions) may be spurious even when it holds necessarily, that is, when it holds in all possible worlds; in such cases, let us say that a *strict logical inconsistency* (or simply *strict inconsistency*) obtains. For example, consider the following material conditional: "If Joe has a picture of a round square, then 2 + 2 = 5." This strange conditional is not only true, but necessarily true: a false antecedent suffices to make a material conditional true, and the antecedent of this conditional is not just false, it is necessarily false. Here, then, we have an example of a material conditional that is necessarily true even though there is no relevance relationship whatsoever (either in syntax or semantics) between the propositions expressed in the antecedent and consequent. Indeed, the statement "Joe has a picture of a round square" is strictly, though trivially, inconsistent with every possible statement.

By contrast, a strict logical incompatibility relation holds only when there is a direct conflict between two (or more) strictly inconsistent statements.[36] For example, the proposition *that John is a bachelor* is strictly logically incompatible with the proposition *that John is married*, for the term 'bachelor' means (among other things) unmarried. By contrast, let us say that *the round-square thesis* states that a round square exists and *the fluffy-kitten thesis* states that

a fluffy kitten exists. The fluffy-kitten thesis is strictly inconsistent with the round-square thesis, but the inconsistency is spurious; the two theses are not strictly (or otherwise) logically incompatible. To put the point another way, the material conditional "If the fluffy-kitten thesis is true, then the round-square thesis is false" is true – indeed, necessarily true – but only trivially so.

3.4 Characterizing Compatibility

For each of the relations discussed above, questions arise about what we should say when we wish to *deny* that two things stand in that relation. For example, any two things which are not incompossible are *compossible*, that is, their co-existence is metaphysically possible. As such, for any two phenomena A and B we might select, the claim "A is incompossible with B" will be semantically equivalent to the claim "It is not the case that A is compossible with B." Likewise, any two propositions which are not inconsistent are *consistent*. Unfortunately, matters are more complicated when it comes to denying that a relevance-relation of incompatibility (whether metaphysical, logical, or conceptual) obtains. Any two things which are metaphysically incompatible are also incompossible, but not vice versa (since incompossibility may be spurious but metaphysical incompatibility cannot be). As such, one may deny that two things are metaphysically incompatible without denying that the things are incompossible (and the same goes, *mutatis mutandis*, for logical incompatibility and logical inconsistency). Now, we could agree to use 'compatibility' very narrowly, such that two things are *compatible* so long as they are not *incompatible*. However, if philosophers were to use 'compatibility' in this narrow way while holding fixed the wording in standard definitions of 'compatibilism,' anyone who categorically rejects causal luck solutions to E2 would qualify as a compatibilist. In other words, this use of 'compatibilism' would open up logical space for compatibilists who are incompossibilists (Mickelson 2012, 2015a) – and, indeed, this space has been claimed by some impossibilists who, by all appearances, choose to identify as "compatibilists" in order to emphasize that they reject incompatibilist (i.e. causal-luck) solutions to E2 of the explanation problem (e.g. Levy 2011, p. 1; see also Mickelson 2019b).

4 Conclusion

The terms 'incompatibilism' and 'compatibilism' were developed for use within the classical analytic paradigm, and they have become substantively ambiguous since free-will theorists started using them outside that dialectical context. It is no longer the case that there are only two basic positions one may take regarding the relationship between free will and determinism, and it may be best for us to stop using language which misleadingly implies otherwise. To repair the discourse, we might introduce new terms which reflect the distinct relations and explanations that are most important in contemporary discussions of the problem of determinism. As suggested previously, we might avoid the term 'incompatibilism' by referring to philosophers who accept an incompossibility solution to the correlation problem and a causal luck solution to E2 as *causal luck incompossibilists* (as opposed to referring to them with the degenerated term 'incompatibilists'), and to causal luck incompossibilists who also endorse a source solution to E1 as *causal-luck source incompossibilists*. By contrast, we might classify incompossibilists who endorse a source solution to E1 and a constitutive luck solution to E2 as *constitutive luck source incompossibilists*, and so on. While I find questions about how to best define and/or redefine certain basic terms of art to be worthy of serious consideration

(e.g. Mickelson 2015a), such matters cannot be settled here. The modest goal of the present chapter has been to clarify the language and dialectical structure of the contemporary debate in a way that leaves the reader with a better understanding of the problem of determinism and its surprising array of rival solutions.

Acknowledgments

Thanks to the University of Gothenburg and the Swedish Research Council for funding this chapter. For their help in developing the ideas presented here and/or comments on early drafts of this chapter, I am deeply indebted to Derk Pereboom, Christian Munthe, Michael McKenna, Robert Hartman, Gunnar Björnsson, Keith Lehrer, Peter van Inwagen, Seth Shabo, Randy Clarke, Kadri Vihvelin, Neil Levy, Maria Sekatskaya, Audun Bengtson, Ragnar Francén Olinder, John Eriksson, Saul Smilansky, Alex Skiles, Anna Sophia Maurin, Christian Lee, Per Milam, Sophia Jeppsson, Matthew Jernberg, Taylor Cyr, Graeme Forbes, V. Alan White, and Joe Campbell.

Notes

1 Peter van Inwagen isolates 1965–1985 as the "classical period" of analytic work on free will; he has also (independently) suggested that the leading free-will theorists in this period were working within a distinctive paradigm (van Inwagen 2017). The terms 'compatibilism' and 'incompatibilism' were coined by Keith Lehrer during this period. To my knowledge, the terms 'compatibilist' and 'incompatibilist' were first used in print in Lehrer's (1960) dissertation, though more standard characterizations of these terms first appeared in print eight years later (Cornman and Lehrer 1968, p. 130). It appears that the corresponding terms 'compatibilism' and 'incompatibilism' were first used in print by van Inwagen in his (1969) dissertation, but van Inwagen credits Lehrer with the coining of these terms as well (in correspondence; see also van Inwagen 2017; 1999, p. 342, n. 2). Lehrer (in correspondence, 2020) has confirmed that he introduced 'incompatibilism' to name a view about the *relevance* of determinism to free will. In saying this, Lehrer means to draw on the familiar conversational notion of *relevance* (as opposed, for example, to some technical notion developed within so-called *relevance logics*), and he is clear that the relevance relation he has in mind is not fully captured by the non-relevance relation of (strict) logical inconsistency. The general notions of *strict logical incompatibility* and *conceptual incompatibility* (Sec. 2) seem to fit well with the notion of relevance that Lehrer has in mind. (The terms 'compossibilism' and 'incompossibilism' were introduced in Mickelson 2012).

2 The classical bipartite (in)compatibilism taxonomy of free-will views was preceded by a pre-classical tripartite taxonomy of *libertarianism, soft determinism*, and *hard determinism*. Hard determinism is the explanatory view that determinism is true and we do not have free will because determinism-related factors preclude free will (e.g. Pereboom 2001, p. 323). Since *free-will denialism* is the non-explanatory view that no normal human in the real world has free will, all hard determinists are denialists. *Libertarians* hold that we have free will (i.e. free-will denialism is false) and, since determinism-related factors destroy free will, it must be that determinism is false. *Soft determinists* reject free-will denialism (like the libertarians) but accept that determinism is true (like the hard determinists) – and, by implication, reject the principled view (endorsed by both hard determinists and libertarians) that determinism-related factors stand in an antagonistic relationship to free will. Classical analytic theorists shifted the focus of the free-will debate to this disagreement about the in-principle relationship between free will and determinism, thereby allowing the debate about the *nature* of free will to proceed without anyone having to take a stand on the truth-value of determinism (at the actual world). In short, each view in the pre-classical tripartite taxonomy took a

stand on three contentious issues: (1) the existence of free will, that is, whether free-will denialism is true, (2) the truth-value of determinism, and (3) whether determinism conflicts with (undermines, precludes, or stands in some broadly antagonistic relation to) free will. Classical analytic theorists aimed to track the third issue with their Compatibility Problem.

3 Notice that James manages to express the idea that naturalistic elements which account for the evolution of the universe make one unique future *inevitable* (by eliminating literally all actual-sequence leeway from the world) without explicit appeals to causation or laws of nature. As such, James shows that it is possible to capture the traditional "literally zero leeway" notion of determinism without having to set foot in the "morass" of causation (see also n. 21 *infra*).

4 Some philosophers assume that 'determinism' should be used to pick out the doctrine that *deterministic* causation and/or *deterministic* laws of nature obtain, where the term 'deterministic' may be defined however one likes. As a result, 'determinism' is now a substantively ambiguous term. For example, some philosophers hold that 'determinism' should be used to name the view that naturalistic factors in the physical universe eliminate all actual-sequence leeway from its evolution, but only *ceteris paribus* (such that the world has, *ceteris paribus*, one unique future) (e.g. Dennett 2003; Sehon 2010). More specifically, one might use 'determinism' to name the thesis that there is no actual-sequence leeway in the world *on the condition that the universe remains causally closed* (e.g. so long as the world is not prematurely destroyed in a collision with another universe in the multiverse or miraculously intervened in by God). The addition of *ceteris paribus* clauses to the Jamesian notion of determinism might seem like an improvement for those who think that free-will theorists should define 'determinism' in a way that reflects the best physics and/or metaphysical theories of the day. In my assessment, the problem of determinism (understood as a means of addressing the problem of free will via the correlation and explanation problems) is best framed using a doctrine which is the limiting case for minimal actual-sequence leeway (i.e. a doctrine which asserts that there is *literally zero* such leeway), for this is the most straightforward way to test whether or not the presence/absence of actual-sequence leeway is *relevant* to such things as "the ability to do otherwise." To repair the discourse, we might introduce a new term (preferably one unrelated to any of the vocabulary terms used by the physicists) such as "inevitabilism" to name the dialectically important doctrine that one unique future is *literally* – and not simply *ceteris paribus* – inevitable given naturalistic factors which hold in the present/past (for discussion, see Mickelson 2019a and Mickelson "Hard Times for Hard Incompatibilism," ms.; see also n. 21 *infra*).

5 The general framework outlined here is not an attempt to persuade readers of what the dialectical structure of the free will debate *should* look like; it aims, more centrally, to be (loosely speaking) a rational reconstruction of what the extant free will debate *does* look like. I believe this framework offers a superior map of the territory on which the free will debate is – and traditionally has been – playing out.

6 This question is typically raised after providing a vignette in which normal humans are performing actions in a world at which determinism is (by stipulation) true. In recent years, those who advocate a "no" answer to the correlation problem have begun using so-called *manipulation arguments* (Mickelson 2017) as intuition pumps for the "no" answer. For example, Alfred Mele's *revised* Zygote Argument (introduced in Mele 2013; see also Mele 2017, 2018, 2019b, 2022) is a (relatively weak) intuition pump for an incompossibilist solution to the correlation problem; unlike the *original* Zygote Argument (Mele 2006, 2008), the revised Zygote Argument does not aim to solve the explanation problem (Mickelson 2015b, 2017, 2019b). Other manipulation arguments, including Pereboom's Four-Case Argument (Pereboom 2001, 2014) and my Master Manipulation Argument (Mickelson 2019b), speak to both the correlation problem and the explanation problem.

7 It has been suggested that one answer E2 by appeal to intuition, for example, by reporting that one has the *intuition* that determinism-related factors undermine free will. I do not find this suggestion compelling. Rather, I think that one's intuitions may lead them to form a *hypothesis* about what precludes free will, but any such hypothesis may be subject to testing which demonstrates that it is mistaken. This is perhaps most evident in the literature on manipulation arguments, which exploit the intuition that a person who is subject to manipulation is unfree even though it is essential

to the success of the argument that, contrary to initial appearances, there is nothing about the manipulation per se that poses a threat to free will. Likewise, abductive manipulation arguments may be developed to test and attack the hypothesis that determinism poses a threat to free will, placing the burden of proof on libertarian theorists to defend that hypothesis against known rivals (Mickelson 2015b, 2017, 2019a, 2019b, 2021a and "Hard Times for Hard Incompatibilism," ms.).

8 The assumption of anthropocentric possibilism is evident, for example, in van Inwagen's "mysterianism" position: he contends that free will is a mystery because he believes that we have free will even though he grants that there is a strong and unanswered case for *anthropocentric impossibilism*, that is, the view that it is metaphysically impossible for a normal human to exercise free will (e.g. van Inwagen 2000; Campbell and Lota CH. 8 this volume). Notably, it has recently become popular to use the term 'impossibilism' exclusively to name the maximally bold impossibilist thesis that it is metaphysically impossible for anyone (i.e. for any metaphysically possible being, however godlike) to exercise free will, and 'possibilism' to pick out the negation of this unqualified impossibilism (e.g. McKenna and Pereboom 2016). I follow the latter convention here, but I warn against conflating impossibilism and species anthropocentric impossibilism. While all arguments for impossibilism are arguments for anthropocentric impossibilism, the converse is not true. For example, Pereboom's arguments for hard incompatibilism imply that anthropocentric impossibilism is true but that impossibilism is false (e.g. Pereboom 2014, pp. 4–6). By contrast, some arguments (e.g. the Basic Argument and some arguments based on the paradox of moral luck) conclude to impossibilism and, *a fortiori*, support anthropocentric impossibilism, but do this by reasoning which implies that hard incompatibilism is false (e.g. Mickelson 2019a, 2019b, and "Hard Times for Hard Incompatibilism," ms.).

9 The classical libertarian solution to E1 narrowly proposes that anthropocentric possibilism is true. However, "broadly libertarian account of free will" is aptly applied to any account of free will which proposes that both incompossibilism and unqualified *possibilism* (the view that it is metaphysically possible for *someone* to exercise free will) are true.

10 One popular test for the adequacy of libertarian solutions to the explanation problem is the so-called "Problem of Enhanced Control," a problem which is typically aimed at *event-causal* (as opposed to *agent-causal*) *libertarianism* (e.g. Franklin 2011). Viewed through the correlation/explanation framework, the problem is roughly this: to be viable, a broadly libertarian response to E1 must identify a necessary condition on free will which cannot be satisfied when determinism is true (i.e. some necessary condition that is missing from all compossibilist accounts of free will), such that it is plausible to think that someone who satisfies that condition would (in virtue of satisfying that condition) have *more control* than would a person who satisfies the necessary conditions on free will named in the best compossibilist accounts of free will. It would be unsatisfactory for aspiring libertarians to respond to E1 by insisting that *indeterminism itself* (the mere negation of determinism) or some more specific indeterminism-related factor (e.g. *probabilistic* causation) is the necessary condition on free will that goes unmet when determinism is true. As such, those who believe that (some species of) indeterminism "helps" a person to have more control than they would have in a world at which determinism is true – namely, those who endorse a libertarian account of free will – owe their audience a story about *how* indeterminism helps. E1 captures and illuminates the dialectical demand for that story.

11 Arguably, the classical libertarian solution to E1 is slightly more narrow than this, since it seems to be that classical analytic theorists agreed that free will requires some degree of *contrastive control* over the future (allowing a person to settle which action is ultimately performed and which are not). Some libertarian accounts of the ability to do otherwise are less demanding, such as Robert Kane's "self-forming actions" account (e.g. Kane 1996, 2009, 2016).

12 Recent debates about the Consequence Argument raise the question of what this argument is an argument for (e.g. Campbell 2007, 2008, 2010; Shabo 2011; Bailey 2013; Sartorio 2016; Capes 2019). These discussions do not take into account that there are two substantively different ways of interpreting van Inwagen's formal statements of the Consequence Argument. First, we may interpret these logic-text proofs – which van Inwagen characterizes as the same argu-

ment done three ways – against the background assumptions of the classical analytic paradigm. In this case, van Inwagen's formal arguments technically conclude to a conditional thesis which is equivalent to classical incompossibilism, but this conclusion entails (via the classical bridge inference) classical incompatibilism. Second, we may untether these formal statements of the Consequence Argument from the background assumptions of the classical analytic paradigm (i.e. the now question-begging classical assumptions which underwrite the classical bridge inference). In this case, the conditional conclusion of van Inwagen's arguments do not entail or suggest any particular solution to E1 or E2 of the explanation problem. When a logic-text statement of the Consequence Argument is removed from its original dialectical context, its conditional conclusion must be taken as it stands, that is, as an inconsistency thesis which provides nothing more than a solution to the correlation problem. This non-classical Consequence Argument – just like Mele's non-classical Zygote Argument (Mickelson 2015b, 2021a, and "Motte-and-Bailey Incompatibilism," ms.) – may be used to motivate a constitutive luck solution to E2 of the explanation problem (which implies that all causal luck solutions are wrong). For further discussion, see my "The Consequence Argument: An Argument for Incompatibilism?," ms.).

13 Frustration with the community's commitment to the classical analytic definition of 'free will' played a role in John Martin Fischer and Mark Ravizza's development of s*emi-compatibilism*, a view about the compossibility of determinism-related factors and the type of control over/ownership of one's actions that is required for *moral responsibility* (e.g. Fischer and Ravizza 1998). According to the semi-compatibilist, determinism-related factors pose no threat to moral responsibility irrespective of whether such factors destroy our ability to otherwise (where this is understood neutrally between the compossibilist and incompossibilist reading). Notably, the latter claim is also accepted by some source compossibilists (e.g. Harry Frankfurt, as well as some anti-compossibilist source impossibilists, for example, G. Strawson, as discussed in main text). When we hold fixed the anachronistic classical definition of 'free will,' so-called semi-compatibilists are committed to the odd-sounding claim that moral responsibility does not require free will.

14 G. Strawson argues that the source condition on free will is impossible for anyone or anything to satisfy, which means that he endorses (unqualified) impossibilism. Pereboom sets his view apart from G. Strawson's in part by insisting that a being with broadly libertarian, "law-overriding" agent-causal powers could (i.e. in some possible world does) satisfy the source condition on free will and act freely (Pereboom 2001, pp. 85–86, 128). Notably, Pereboom also accepts free-will denialism, the view that *we* (i.e. members of the class *normal humans*) do not have free will, by reasoning which seems to commit Pereboom to anthropocentric impossibilism (for full discussion, see Mickelson "Hard Times for Hard Incompatibilism," ms.)

15 It should now be clear that the incompossibilist, when faced with the question "Why can't a normal person act freely assuming determinism is true?," cannot adequately answer the challenge by saying "Well, because the actions of such a person would be *completely settled by factors beyond his control*" (e.g. Capes 2019). The challenge posed by E2 is the problem of *pinpointing* the factor(s) which make a person unfree when determinism is true. At best, an allusion to unidentified freedom-undermining "factors" indicates that incompossibilism is not a *brute fact* – but this simply confirms the legitimacy of E2 without providing a solution to it.

16 This interpretation stands in sharp contrast to the standard narrative within the moral luck literature, according to which the traditional problem of determinism is equivalent to what moral-luck theorists would call the problem of antecedent causal luck, where "causal luck" is understood narrowly (and, I think, mistakenly) as *deterministic* antecedent causal luck (e.g. Latus 2000, p. 167, n. 6). For full discussion, see Mickelson 2019b.

17 Notably, moral-luck theorists use the term 'control' (e.g. in statements of the "control principle" generates the paradox of moral luck) to pick out the same type of basic-desert-grounding control that most free-will theorists now pick out with the term 'free will' (Mickelson 2019b).

18 I say "inspired by" because Nagel's original taxonomy of luck/factors beyond our control was unprincipled: Nagel's categories were neither mutually exclusive nor collectively exhaustive (he did not suggest they were) in ways that make it unsatisfactory for use in a discussion of rival solutions

to E2 (Mickelson 2019b). Unfortunately, there is no way to capture the overlapping metaphysical problems in the literatures on free will, moral luck, and constitutive luck without making some adjustments to (or completing replacing) the distinctive jargon used in each. For those who resist updating Nagel's terms/categories vis-a-vis moral luck, they are welcome to apply any labels they prefer.

19 I understand constitutive luck broadly to include both a lack of control over actual and modal facts about a person. On this understanding, we may distinguish the type of constitutive luck one suffers from when the endowment one has in the actual world is beyond one's control ("endowment luck") and the type one suffers when their essential properties, roughly the properties that underwrite one's cross-world identity, if there are any such properties. That is, it might be that one's entire "modal profile" (roughly one's every possible endowment) is beyond one's control ("modal profile luck"). Following this line of thought, the problem of free will quickly becomes a debate about the metaphysics of personal identity rather than causation or the laws of nature.

20 By endorsing James's definition of 'determinism', I have in effect defined 'determinism' to denote the thesis that the naturalistic factors which account for the evolution of the physical universe (laws, causation, or the like) are *strong*, *deterministic*, and *unconditional*; philosophers who think we should add some kind of *ceteris paribus* clause to the traditional (Jamesian-style) statements of determinism are, in effect, using 'determinism' to pick out the slightly different thesis that the relevant evolutionary factors are strong, deterministic, and *conditional* (Mickelson 2019a and Mickelson "Hard Times for Hard Incompatibilism" ms.; see also n. 4 *supra*). Given the latter definition, determinism may be true even in a world at which God regularly intervenes in the natural order – which is interesting because it means that a libertarian might argue that God's intervention introduces freedom-relevant actual-sequence leeway into the world even though "determinism" is true (e.g. Stone 1998). The fact that the latter definition of 'determinism' blurs the line between compossibilists and broadly libertarian accounts of free will speaks against its use.

21 Notably, some philosophers have advanced a doctrine paradoxically called "Humean determinism" (Beebee and Mele 2002). I say "paradoxically" because so-called Humean determinism may be true in a world with weak laws of the sort the permit unlimited moment-to-moment actual-sequence leeway in the world (as depicted in van Inwagen's forking-paths diagram earlier). This means that the phrase "Humean determinism" refers to a paradigmatic doctrine of *indeterminism* by any traditional definition of the term 'determinism'. Since this is an odd development, the reader may wonder how such a twist of terminology might have arisen. In brief, the term "Humean determinism" was introduced when it was noted that a certain logical entailment thesis that van Inwagen had referred to by the name 'determinism' (because the entailment thesis served, in logic-text proofs, as a proxy for the metaphysical doctrine depicted in his no-forking-paths diagram) was open to a broadly Humean interpretation. Early commentators took the Humean interpretation of *the proxy thesis* as evidence that there is a Humean interpretation of *traditional zero-leeway determinism*. However, the availability of a Humean interpretation of van Inwagen's *proxy thesis* is better taken as evidence that this proxy failed to capture its target. On the latter line of reasoning, van Inwagen extended the term 'determinism' to his proxy thesis *by mistake*, and it is upon this mistake that the unhappy notion of "Humean determinism" took root. (For a full discussion, see Mickelson "Humean-law Determinism, Humean-law Compatibilism, and The Consequence Argument," ms.)

22 Standard abductive reasoning supports the following inference: If (contrary to what libertarian accounts of free will propose) indeterminism doesn't "help" a person to act freely, then (contrary to what incompatibilists claim) determinism doesn't "hurt" – and, indeed, whether determinism is true or false is *irrelevant* to whether we have free will (Mickelson 2019a, 2019b, and "Hard Times for Hard Incompatibilism", ms.).

23 Thomas Nagel expresses similar views in his discussions of the paradox of moral luck (Mickelson 2019b).

24 Carolina Sartorio comes close to endorsing a strict circumstantial luck solution to E2 when she suggests that determinism would pose no threat to free will but for the fact that we are "causally impotent" vis-à-vis the events in the remote past when determinism is true, but she might

also be interpreted as endorsing a causal/circumstantial hybrid solution (which is distinctive because of her unusual emphasizes the freedom-undermining role played by our lack of control over circumstances in the remote past) (e.g. Sartorio 2016, pp. 151–152). We might classify those incompossibilists who contend that past circumstances (apart from such things as the laws of nature and one's constitutive properties) pose a distinctive threat to free will as *circumstantial luck incompossibilists*.

25 Philosophers who lack a firm grasp on this in-house debate regarding the solution to E2 often assume that an incompossibilist may accept the conclusion of every argument against compossibilism, but this is a mistake. For example, Gregg Caruso seems to endorse both Pereboom's "hard incompatibilism" and Levy's "hard luck" impossibilism (e.g. Caruso 2019; Dennett and Caruso 2021, p. 176; see also Mickelson 2021b), but this position (taken at face value) is untenable. While both Pereboom and Levy are incompossibilists, they endorse rival solutions to E2 of the explanation problem. Levy – as Caruso notes (Dennett and Caruso 2021, p. 196) – self-identifies as a *compatibilist* to indicate that he rejects the "incompatibilist" thesis that deterministic laws/causation poses a distinct threat to free will (Levy 2011, p. 1, n.1). In more generic language, one cannot accept both (i) Pereboom's incompatibilism (according to which the truth of determinism is negatively, though *not spuriously*, correlated with the existence of free human agents), and (ii) Levy's anti-incompatibilist view (which I consider a fleshing out of G. Strawson's position, e.g. Mickelson 2019b) that determinism is negatively, but only *spuriously*, correlated with the existence of free human agents.

26 The notion of sourcehood which drives G. Strawson's constitutive-luck source impossibilism also fits well with Nagel's observation that many people are paradoxically committed to the existence and impossibility of moral luck (Mickelson 2019b). It is not at all clear, however, that Pereboom's broadly libertarian characterization of the source condition fits with his preferred incompatibilist (a.k.a. causal luck incompossibilist) solution to E2 (Mickelson "Hard Times for Hard Incompatibilism," ms.).

27 In a recent "glossary for the uninitiated," Alfred Mele asserts that he is "following standard practice" when he defines 'incompatibilism' to pick out a perhaps spurious incompossibility claim (e.g. Mele 2019a, p. 1, n.1; see also Mele Ch. 31 this volume); indeed, he asserts that those who use the term in any other way are using it in a "nontraditional" (Mele 2017, p. 6, n. 4) or "nonstandard" way for which he has "never had any use" (Mele 2019b, p. 3, n. 1). Mele's failure to provide evidence for these claims is notable, for the claims appear to be false. Philosophers regularly use 'incompatibilism' (following Lehrer) to pick out a type of explanatory-relevance view, that is, a species of causal luck incompossibilism, as opposed to the relatively modest, non-explanatory thesis of mere incompossibilism. Given such facts, it is puzzling that Mele has singled out only two philosophers by name – Mickelson (Mele 2017, p. 6, n. 4) and Levy (Mele Ch. 31 this volume) – as examples of philosophers who use the term 'incompatibilism' differently than he does, that is, to refer to something other than mere incompossibilism. After all, Mele's comments about 'incompatibilism' commit him to the view that Keith Lehrer (who coined the term), Derk Pereboom (2001, 2014), Randy Clarke (2003), Carolina Sartorio (2016), Kadri Vihvelin (2008, 2013), Michael McKenna (2010), John Martin Fischer and Mark Ravizza (1998, p. 151) and many other leading figures in the debate who regularly use the term 'incompatibilism' to denote causal luck incompossibilism are guilty of using that term in a nonstandard/nontraditional way. Moreover, Mele's claims about his own use of the term 'incompatibilism' are also problematic. Mele used the term 'incompatibilism' to pick out a species of causal luck incompossibilism when he claimed that Pereboom's Four-Case Argument "fails as an argument for incompatibilism" (Mele 2005, p. 80; 2006, pp. 144, 189; see also Mele 2008, p. 278). Mele's attack on Pereboom's best-explanation reasoning narrowly targets the explanation step of Pereboom's argument, leaving its counterexample step and generalization step (which supports an incompossibilist solution to the correlation problem) untouched (Mickelson 2017). Mele's treatment of the term 'incompatibilism' has also contributed to the common misconception that Mele's revised Zygote Argument aims to support the same explanatory conclusion as its predecessor, the Four-Case Argument (e.g. Mele 2005, 2006, 2008,

2013, 2017, 2018, 2019b, 2022); in fact, the former concludes to mere incompossibilism but the latter concludes to causal luck incompossibilism. Notably, Mele has introduced no new terms which would allow him (or his interlocutors) to easily distinguish between the relatively modest correlation-claim conclusion of his *revised* Zygote Argument (Mele 2013, 2017, 2018, 2019b, 2022) and the familiar explanatory conclusion of Pereboom's Four-Case Argument (see Mickelson 2015b, 2017, and 2021a for discussion); since both aim to support *at least* incompossibilism, Mele classifies both as "arguments for incompatibilism" and leaves it at that. Finally, because Mele has used different definitions of the term 'incompatibilism' in different dialectical contexts – for example, when he is advancing a criticism of Pereboom's argument (Mele 2005) versus when he is responding to a structurally identical criticism of his Zygote Argument (e.g. Mele 2013; Mickelson 2015b, 2021a) – Mele is vulnerable to the charge of sophistical *motte-and-baileying* on the term 'incompatibilism' and its cognates (Mickelson "Motte-and-Bailey Incompatibilism," ms.).

28 For an interesting discussion of ongoing disputes about the demarcations between these relations, see Wirling 2020.

29 I have been told (in public discussion and private correspondence) that most of the words/phrases listed here are ambiguous in natural language, such that one may use most of the phrases listed here without indicating that an antagonistic relevance relation holds between things. I doubt this is correct, but I will not challenge the point here. I raise the issue only to make the reader aware of the fact that at least some mainstream analytic philosophers do not use these terms to indicate the presence of a relevance relation.

30 In this chapter, I use the term 'proposition' to pick out abstract truth-value-bearing entities. While I recognize the distinction between propositions and statements, I will treat the distinction loosely here because two (or more) statements may stand in any of the logical relations described here, and likewise two (or more) propositions may stand in any of the logical relations described here.

31 I purposely avoid language which suggests that every relevance relation is a type of correlation relation (e.g. I avoid phrases such as 'relevance correlation'). I think this way of speaking would exacerbate current confusion by needlessly blurring the sharp distinction between perhaps spurious (i.e. perhaps *non-relevance*) correlation relations and (always, by definition, non-spurious) relevance relations.

32 One may draw upon the recent "Dependence Solution" to the problem of free will and foreknowledge to flesh out the asymmetrical dependence relationship of God's foreknowledge on our actions, such that God's knowledge of the future is not a threat to free will (e.g. Hunt and Zagzebski 2022) – even if, as I have stipulated here, something else (determinism-related causal factors beyond one's control) in the world is "doing the work" of destroying free will. (Notably, use of terms 'incompatibility' and 'incompatibilism' has spread to the free will/foreknowledge literature, where the terms are problematically ambiguous in the way that they are in the free will/determinism literature.)

33 I do not find it helpful to use the term 'incompossibility' to refer to the strict logical inconsistency relation; for those who do, I strongly suggest using the qualifier 'logical'. Failing to do so may lead to confusion, for example, between questions/views about how the *truth-values* of certain propositions are related (if at all) across possible worlds, and questions/views about the possible *co-existence* of a specific proposition (i.e. some truth-value-bearing entity) and some other thing (e.g. another proposition, some other type of abstract object, some physical object, etc.) should be carefully distinguished from questions. For example, to say "Determinism is *logically incompossible* (a.k.a. strictly logically inconsistent with) the thesis that someone has free will" is not equivalent to saying "Determinism is *metaphysically incompossible* with free will." The latter is a claim about the possible co-existence of the thing picked out by the term 'free will' (which does not have the ontological status of a proposition) and the doctrine of determinism (which has the ontological status of a proposition); this metaphysical incompossibility claim says nothing about the *truth-value* of determinism. The former claim, by contrast, assumes the existence of two propositions and asserts something about the truth-values of two propositions. Notably, one may easily reject the latter (metaphysical incompossibility) claim without taking a stand on the former (strict logi-

cal inconsistency/logical incompossibility) claim – and, indeed, without taking a stand on any of the central questions in the free-will debate.

34 Restated in the language of possible worlds, we may say that statements are logically inconsistent if and only if they cannot be true at the same time *at the actual world* (irrespective of the reason/reasons why). For examples of this generic, non-relevance definition of 'inconsistency', see, for example, Bergmann et al. 1990, pp. 2, 16; Tidman and Kahane 2003, p. 16. Given this definition of 'logical inconsistency,' both contrary statements (i.e. those which can both be false but cannot both be true) and contradictory statements (i.e. those which cannot both be true and cannot both be false) qualify as logically inconsistent statements – even when the contrariety relation is rooted in the semantic content of the two statements and not in their logical form, for example, "The Taj Mahal is pink all over" and "The Taj Mahal is blue all over" (Layman 1999, p. 144). As such, I consider definitions of 'logic inconsistency' on which two propositions/statements are inconsistent if and only if their conjunction entails a contradiction (e.g. Barker 1985, p. 348) to be overly narrow for general purposes, even if there are contexts in which this type of stipulative/technical definition is useful.It is worth noting that introductory logic texts which focus primarily on propositional/sentential logic often describe "truth-table tests" for inconsistency and narrowly define 'inconsistency' in terms of such tests (e.g. Lemon 1992, p. 69; Hurley 1994, pp. 322–323, 327–328; Hurley and Watson 2018, pp. 360, 724). The truth-table method is designed to identify a conflict in the respective logical forms of two (or more) statements, where this syntactic conflict alone ensures that the statements cannot be true at the same time (i.e. there is no possible assignment of truth-values to the components of the statements on which the lines under the main operator of each statement has the truth-value true). While such truth-table tests are sometimes said to reveal "truth-functional inconsistency" (Bergmann et al. 1990, pp. 73–75), it seems preferable to say that such tests identify a *syntactic logical incompatibility* between two (or more) statements, for the same reason it is preferable to avoid speaking of a "causal correlation" when one means to say that a given correlation is due to a direct causal relation between the correlates (i.e. we have reason to avoid treating causal relations and relevance relations more generally as special types of correlation relations). For present purposes, I set aside uses of the term 'inconsistency' in metatheory (see, e.g., Lemon 1992, pp. 68, 75).

35 Since the material conditional ($p \rightarrow \sim q$) is true so long as its antecedent has the truth-value *false* and/or its consequent has the truth-value *true*, it is also logically equivalent to the disjunction ($\sim p \vee \sim q$).

36 The term 'logical incompatibility' is not a technical term in classical logic – indeed, this term rarely mentioned, let alone defined as a technical term, in any standard classical logic text. Presumably this is because the standard tools of classical logical cannot track relevance relations. I take it that recent work on relevance logics (Mares 2020), grounding relations between propositions (e.g. Fine 2012a; 2012b), and the logic of 'because' (e.g. Schnieder 2011) is aimed (in part) at fleshing out the sorts of logical relevance relations which fall under (what I am calling) logical incompatibility and strict logical incompatibility.

Bibliography

Anderson, M.B. (2019). Moral luck as moral lack of control. *Southern Journal of Philosophy* 57 (1): 5–29.

Ayer, A.J. (1954/2009). Freedom and necessity. In: *Free Will*, 2e (ed. D. Pereboom), Oxford: Oxford University Press. 139–147.

Bailey, A. (2013). Incompatibilism and the past. *Philosophy and Phenomenological Research* 85 (2): 351–376.

Baker, L.R. (2006). Moral responsibility without libertarianism. *Noûs* 40 (2): 307–330.

Barker, S. (1985). *The Elements of Logic*, 4e. New York: McGraw-Hill Book Company.

Beebee, H. and Mele, A. (2002). Humean compatibilism. *Mind* 111 (442): 201–223.

Bergmann, M., Moor, J., and Nelson, J. (1990). *The Logic Book*, 2e. New York: McGraw-Hill Publishing Company.

Bobzien, S. (1998). *Determinism and Freedom in Stoic Philosophy*. Oxford: The Clarendon Press; Oxford University Press.

Campbell, J. (2007). Free will and the necessity of the past. *Analysis* 67 (2): 105–111.

Campbell, J. (2008). Reply to Brueckner. *Analysis* 68 (3): 264–269.

Campbell, J. (2010). Incompatibilism and fatalism: reply to Loss. *Analysis* 70: 71–76.

Campbell, J. (2011). *Free Will*. Massachusetts: Polity Press.

Capes, J. (2019). What the consequence argument is an argument for. *Thought: A Journal of Philosophy* 8 (1): 50–56.

Caruso, G. (2019). A defence of the luck pincer: why luck (still) undermines moral responsibility. *Journal of Information Ethics* 28 (1): 51–72.

Chisholm, R. (1964/2009). Human freedom and the self? In: *Free Will*, 2e (ed. D. Pereboom), 172–195. Hackett Publishing Company.

Clarke, R. (2003). *Libertarian Accounts of Free Will*. Oxford: Oxford University Press.

Cornman, J. and Lehrer, K. (1968). *Philosophical Problems and Arguments: An Introduction*. New York: Macmillan Company.

Dennett, D. (2003). *Freedom Evolves*. London: Penguin Books.

Dennett, D. and Caruso, G. (2021). *Just Deserts: Debating Free Will*. Cambridge: Polity Press.

Fine, K. (2012a). Guide to ground. In: *Metaphysical Grounding* (ed. F. Correia and B. Schnieder), 37–80. Cambridge: Cambridge University Press.

Fine, K. (2012b). The pure logic of ground. *Review of Symbolic Logic* 5 (1): 1–25.

Fischer, J.M. (1994). *The Metaphysics of Free Will: An Essay on Control*. Cambridge: Blackwell Publishers.

Fischer, J.M. and Ravizza, M. (1998). *Responsibility and Control: A Theory of Moral Responsibility (Cambridge Studies in Philosophy and Law)*. Cambridge: Cambridge University Press.

Franklin, C.E. (2011). The problem of enhanced control. *Australasian Journal of Philosophy* 89 (4): 687–706.

Griffith, M. (2017). Major positions in the free will debate. In: *The Routledge Companion to Free Will* (ed. K. Timpe, M. Griffith and N. Levy), 1–3. New York: Routledge.

Haji, I. and Cuypers, S. (2006). Hard- and soft-line responses to Pereboom's four-case manipulation argument. *Acta Analytica* 21: 19–35.

Hartman, R.J. (2017). *In Defense of Moral Luck: Why Luck Often Affects Praiseworthiness and Blameworthiness*. New York: Routledge University Press.

Hermes, C. (2013). Truthmakers and the direct argument. *Philosophical Studies* 167 (2): 401–418.

Hermes, C. and Campbell, J. (2012). More trouble for direct source incompatibilism: reply to Yang. *Acta Analytica* 27 (3): 335–344.

Hunt, D. and Zagzebski, L. (2022). Foreknowledge and free will. In: *The Stanford Encyclopedia of Philosophy*, (Summer 2022 Edition) (ed. E.N. Zalta). https://plato.stanford.edu/archives/sum2022/entries/free-will-foreknowledge. (Accessed on 22nd September 2022).

Hurley, P. (1994). *A Concise Introduction to Logic*, 5e. Belmont, CA: Wadsworth Publishing Company.

Hurley, P. and Watson, L. (2018). *A Concise Introduction to Logic*, 13e. Belmont, CA: Cengage Learning, Inc.

James, W. (1884/1897). The Dilemma of determinism. An address to the Harvard divinity students, published in the Unitarian review for September 1884. *Republished in: The Will to Believe and Other Essays (1897)*. London: Longmans, Green, and Co.

Kane, R. (1996). *The Significance of Free Will*. Oxford: Oxford University Press.

Kane, R. (1999). Responsibility, luck, and chance: reflections on free will and indeterminism. *The Journal of Philosophy* 96 (5): 217–240.

Kane, R. (2002). Introduction: the contours of the contemporary free will debates. In: *The Oxford Handbook of Free Will* (ed. R. Kane), 1–41. Oxford: Oxford University Press.

Kane, R. (2004). Agency, responsibility, and indeterminism: reflections on libertarian theories of free will. In: *Freedom and Determinism* (ed. J., Campbell, M. O'Rourke and D. Shier), 70–88. Cambridge: MIT Press.

Kane, R. (2009). Free will and the dialectic of selfhood: can one make sense of a traditional free will requiring ultimate responsibility? *Ideas y Valores* 58 (141): 25–43.

Kane, R. (2016). On the role of indeterminism in libertarian free will. *Philosophical Explorations* 19 (1): 2–16.

Kane, R. (2019). The complex tapestry of free will: striving will, indeterminism and volitional streams. *Synthese* 196 (1): 145–160.

Latus, A. (2000). Moral and epistemic luck. *Journal of Philosophical Research* 25 (1): 149–172. Available online: https://www.iep.utm.edu/moralluc ISSN 2161-0002.

Latus, A. (2001). Moral luck. In: *The Internet Encyclopedia of Philosophy* (ed. J. Feiser). ISSN 2161-0002, https://iep.utm.edu/moralluc/. (Accessed in September 2022).

Layman, C.S. (1999). *The Power of Logic*. Mountain View, CA: Mayfield.

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

Lemon, E.J. (1992). *Beginning Logic*. Indianapolis: Hackett Publishing.

Levy, N. (2011). *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press.

Mares, E. (2020). Relevance logic. In: *The Stanford Encyclopedia of Philosophy* (Winter 2020 Edition), (ed. E.N. Zalta), https://plato.stanford.edu/cgi-bin/encyclopedia/archinfo.cgi?entry=logic-relevance. (Accessed in May 2021).

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

McKenna, M. and Pereboom, D. (2016). *Free Will: A Contemporary Introduction*. New York: Routledge.

McPherson, T. (2021). Supervenience in ethics. In: *The Stanford Encyclopedia of Philosophy*, (Summer 2021 Edition) (ed. E.N. Zalta), https://plato.stanford.edu/archives/sum2021/entries/supervenience-ethics. (Accessed in September 2022).

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*. Oxford: Oxford University Press.

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

Mele, A. (2012). Another scientific threat to free will? *The Monist* 95 (3) Neuroethics: 422–440.

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

Mele, A. (2017). *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*. Oxford: Oxford University Press.

Mele, A. (2018). Diana and Ernie return: on Carolina Sartorio's. *Causation and Free Will. Philosophical Studies* 175 (6): 1525–1533.

Mele, A. (2019a). Free will and moral responsibility: manipulation, luck, and agents' histories. *Midwest Studies In Philosophy* 43 (1): 75–92.

Mele, A. (2019b). *Manipulated Agents: A Window to Moral Responsibility*. Oxford University Press.

Mele, A. (2022). *Free Will: An Opinionated Guide*. Oxford: Oxford University Press.

Mickelson, K.M. (a.k.a. Kristin Demetriou) (2012). Free Will Fundamentals: Agency, Determinism, and (In)compatibility. Dissertation, University of Colorado, Boulder. https://scholar.colorado.edu/concern/graduate_thesis_or_dissertations/g732d9110.

Mickelson, K.M. (2015a). A critique of Vihvelin's three-fold classification. *Canadian Journal of Philosophy* 45 (1): 85–99.

Mickelson, K.M. (a.k.a. Kristin Demetriou) (2015b). The zygote argument is invalid—now what? *Philosophical Studies* 172 (11): 2911–2929.

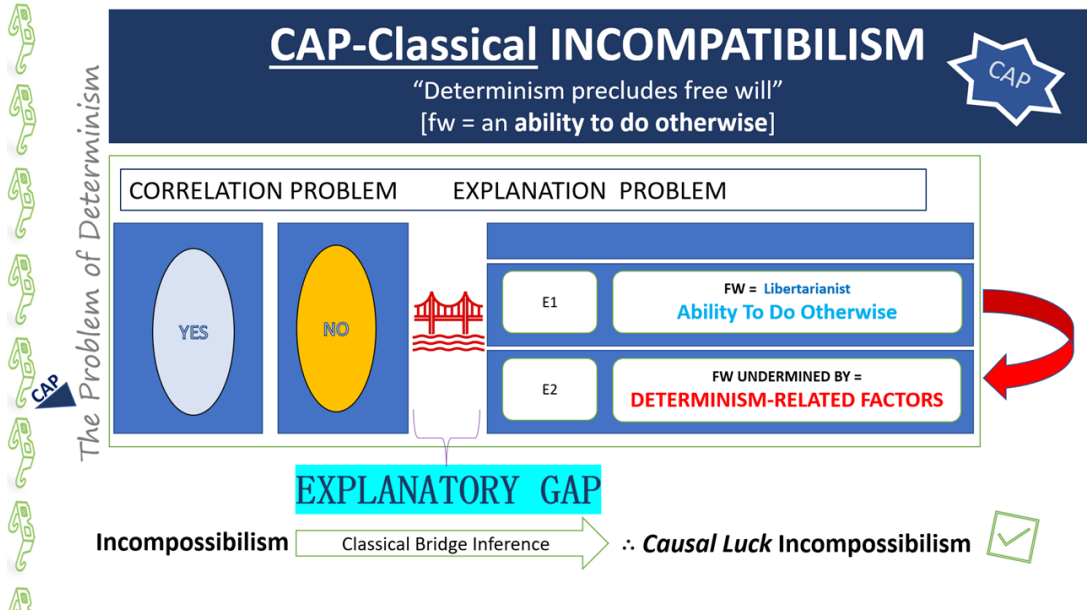
Mickelson, K.M. (a.k.a. Kristin Demetriou) (2017). The manipulation argument. In: *The Routledge Companion to Free Will* (ed. K. Timpe, M. Griffith and N. Levy), New York: Routledge.

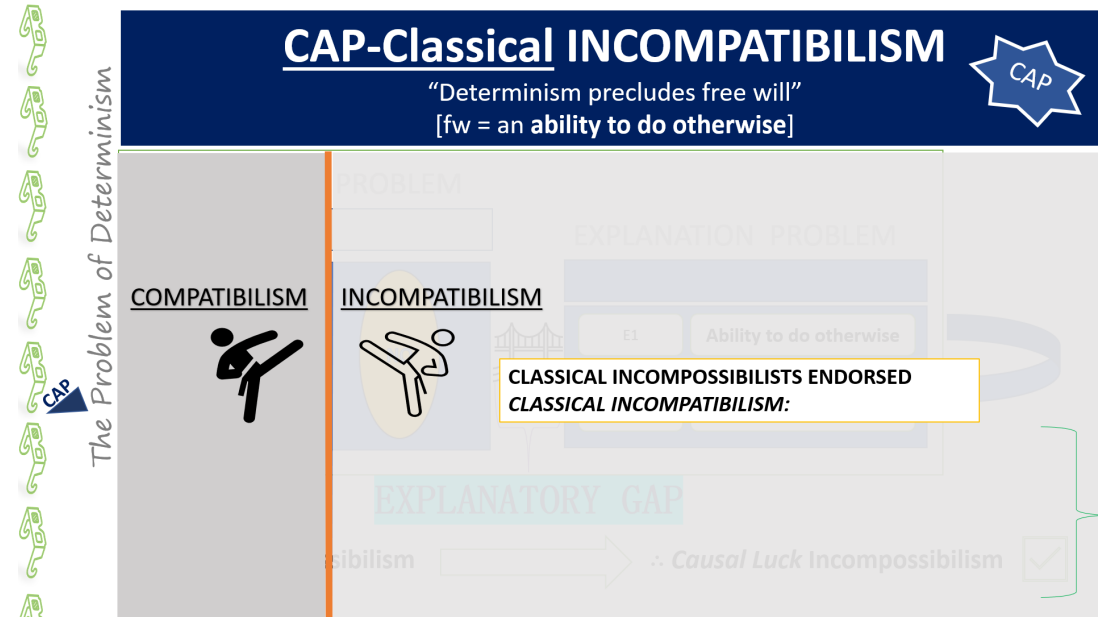
Mickelson, K.M. (a.k.a. Kristin Demetriou) (2019a). Free will, self-creation, and the paradox of moral luck. *Midwest Studies in Philosophy* 43 (1): 224–256.

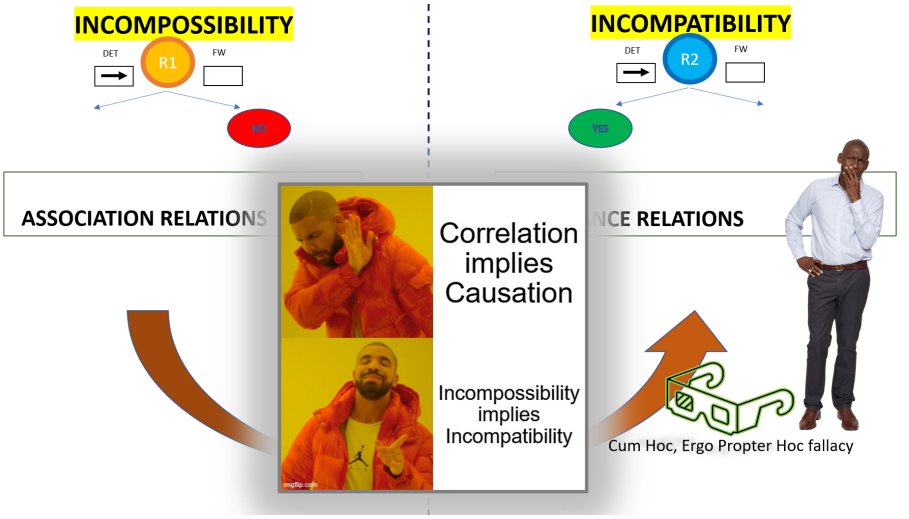
Mickelson, K.M. (a.k.a. Kristin Demetriou) (2019b). The problem of free will and determinism: an abductive approach. *Social Philosophy & Policy* 36 (1): 154–172.

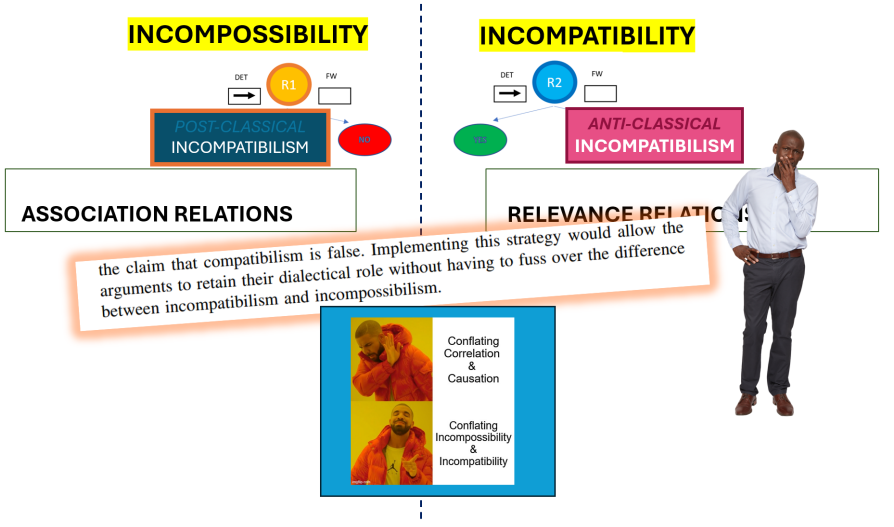
Mickelson, K.M. (a.k.a. Kristin Demetriou) (2021a). The zygote argument is still invalid: so what? Philosophia 49 (2): 705–722.

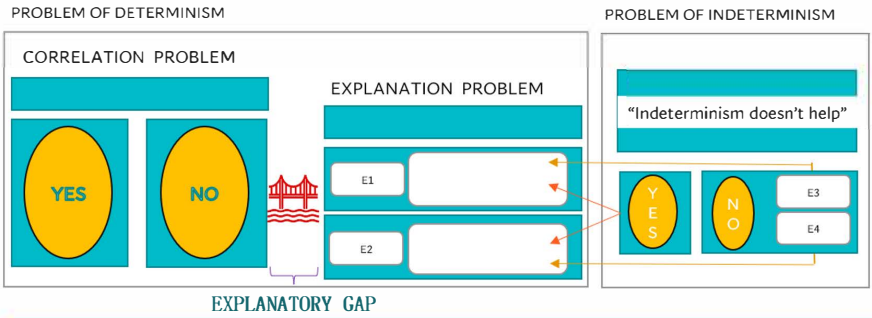
Mickelson, K.M. (a.k.a. Kristin Demetriou) (2021b). Just deserts: debating free will. International Journal of Philosophical Studies 29 (3): 408–412.
Nagel, T. (1976). Moral Luck II. *Proceedings of the Aristotelian Society, Supplementary Volumes* 50: 137–151.
Nagel, T. (1979). *Mortal Questions*. New York: Cambridge University Press.
Nagel, T. (1986). *The View from Nowhere*. Oxford: Oxford University Press.
Nahmias, E., Morris, S., Nadelhoffer, T., and Turner, J. (2006). Is incompatibilism intuitive? *Philosophy and Phenomenological Research* 73 (1): 28–53.
Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.
Pereboom, D. (2005). Defending hard incompatibilism. *Midwest Studies in Philosophy* 29 (1): 228–247.
Pereboom, D. (2009). The stoics. In: *Free Will*, 2e (ed. D. Pereboom), 5–16. Indianapolis: Hackett Publishing Company, Inc. Originally published in *Hellenistic Philosophy* (ed. and trans. B. Inwood and L.P. Gerson). Indianapolis: Hackett Publishing Company (1988).
Pereboom, D. (2014). *Free Will, Agency, and Meaning in Life*. Oxford: Oxford University Press.
Perry, J. (2004). Compatibilist options. In: *Freedom and Determinism* (ed. J. Campbell, M. O'Rourke and D. Shier), 231–254. Cambridge: MIT Press.
Sartorio, C. (2016). *Causation and Free Will*. Oxford: Oxford University Press.
Schnieder, B. (2011). A logic for 'because'. *The Review of Symbolic Logic* 4 (3): 445–465.
Sehon, S. (2010). A flawed conception of determinism in the consequence argument. *Analysis* 71 (1): 30–38.
Shabo, S. (2011). What must a proof of incompatibilism prove? *Philosophical Studies* 154 (3): 361–371.
Statman, D. (2019). The definition of luck and the problem of moral luck. In: *The Routledge Handbook of the Philosophy and Psychology of Luck* (ed. I.M. Church and R.J. Hartman), 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. (1986). *Freedom and Belief*. Oxford: Clarendon Press.
Strawson, G. (1994). The impossibility of moral responsibility. *Philosophical Studies* 75 (1–2): 5–24.
Strawson, G. (2002). The bounds of freedom. In: *The Oxford Handbook of Free Will* (ed. R. Kane), 441–460. Oxford: Oxford University Press.
Strawson, G. (2008). The impossibility of ultimate moral responsibility. In: *Free Will*, 2e (ed. D. Pereboom), 289–306. Oxford: Oxford University Press.
Strawson, G. (2011). Free Will. In: *Routledge Encyclopedia of Philosophy* (ed. E. Craig) London: Routledge. doi: 10.4324/9780415249126-V014-2. https://www.rep.routledge.com/articles/thematic/free-will/v-2 (Accessed in January 2022).
Tidman, P. and Kahane, H. (2003). *Logic and Philosophy*, 9e. Belmont, CA: Wadsworth/Thomson Learning.
van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Oxford University Press.
van Inwagen, P. (1990). Logic and the free will problem. *Social Theory and Practice* 16 (3): 277–290.
van Inwagen, P. (1999). Moral responsibility, determinism, and the ability to do otherwise. *Journal of Ethics* 3 (4): 343–351.
van Inwagen, P. (2000). Free will remains a mystery: the eighth philosophical perspectives lecture. *Philosophical Perspectives* 14: 1–19.
van Inwagen, P. (2017). The problem of fr** w*ll. In: *Thinking about Free Will* (ed. P. van Inwagen), 192–209. Cambridge: Cambridge University Press.
Vihvelin, K. (2008). Compatibilism, incompatibilism, and impossibilism. In: *Contemporary Debates in Metaphysics* (ed. T. Sider, J. Hawthorne and D.W. Zimmerman), 303–318. Malden, MA: Blackwell Publishing.
Vihvelin, K. (2013). *Causes, Laws, and Free Will: Why Determinism Doesn't Matter*. Oxford: Oxford University Press.
Vihvelin, K. (2022). Arguments for incompatibilism. In: *The Stanford Encyclopedia of Philosophy*, (Fall 2022 Edition) (ed. E.N. Zalta and U. Nodelman), https://plato.stanford.edu/archives/fall2022/entries/incompatibilism-arguments/. (Accessed in September 2022).
Wirling, Y.S. (2020). Is backing grounding? *Ratio* 33 (3): 129–137.


CAP-Classical INCOMPATIBILISM
"Determinism precludes free will"
[fw = an ability to do otherwise]
CAP
The Problem of Determinism
CAP
CORRELATION PROBLEM
EXPLANATION PROBLEM
YES
NO
E1
FW = Libertarianist
Ability To Do Otherwise
E2
FW UNDERMINED BY =
DETERMINISM-RELATED FACTORS
EXPLANATORY GAP
Incompossibilism
Classical Bridge Inference
∴ Causal Luck Incompossibilism

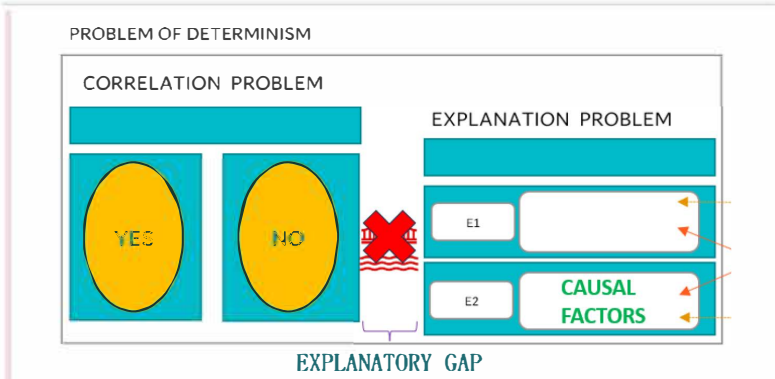

CAP-Classical INCOMPATIBILISM
"Determinism precludes free will"
[fw = an ability to do otherwise]
CAP
The Problem of Determinism
CAP
COMPATIBILISM
INCOMPATIBILISM
CLASSICAL INCOMPOSSIBILISTS ENDORSED
CLASSICAL INCOMPATIBILISM:

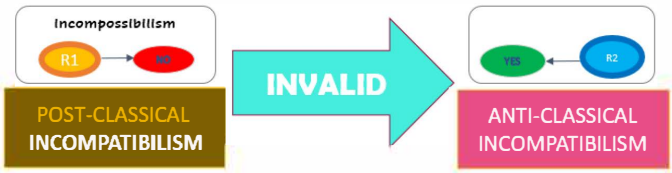

INCOMPOSSIBILITY
DET
R1
FW
INCOMPATIBILITY
DET
R2
FW
ASSOCIATION RELATIONS
Correlation implies Causation
Incompossibility implies Incompatibility
Cum Hoc, Ergo Propter Hoc fallacy

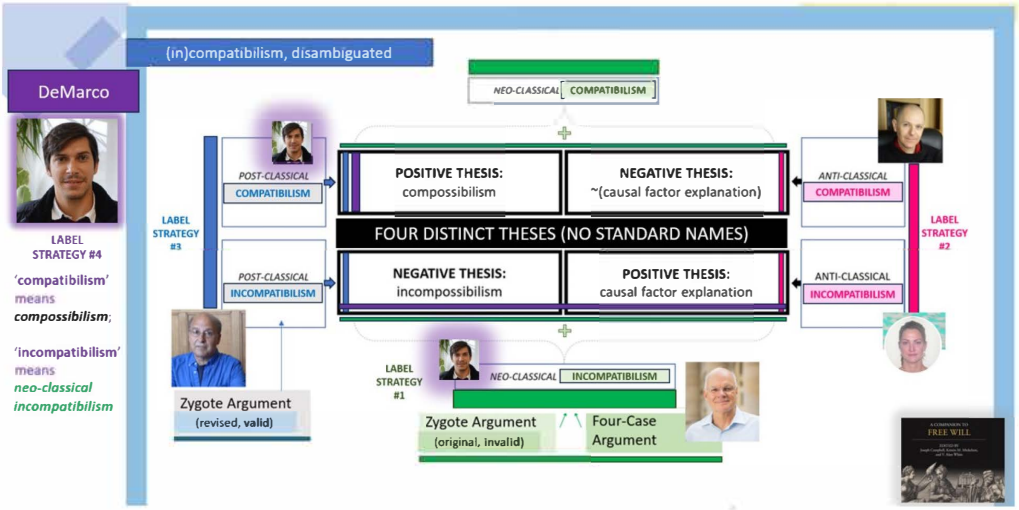

INCOMPOSSIBILITY
DET
R1
FW
POST-CLASSICAL
INCOMPATIBILISM
INCOMPATIBILITY
DET
R2
FW
ANTI-CLASSICAL
INCOMPATIBILISM
ASSOCIATION RELATIONS
RELEVANCE RELATIONS
the claim that compatibilism is false. Implementing this strategy would allow the arguments to retain their dialectical role without having to fuss over the difference between incompatibilism and incompossibilism.
Conflating Correlation & Causation
Conflating Incompossibility & Incompatibility

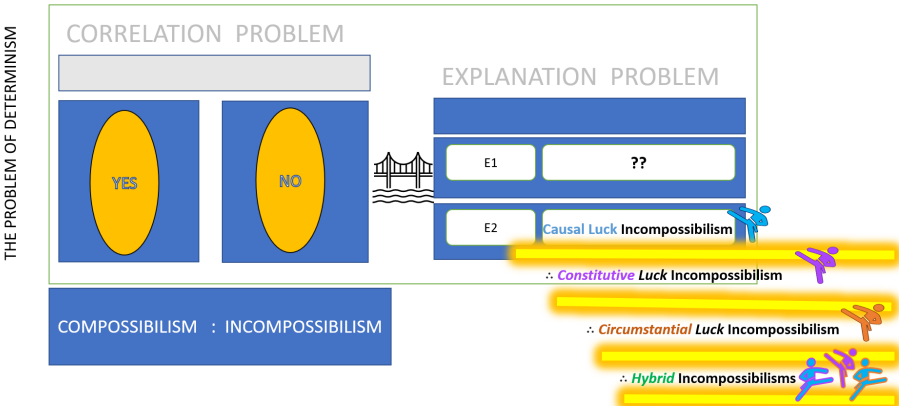


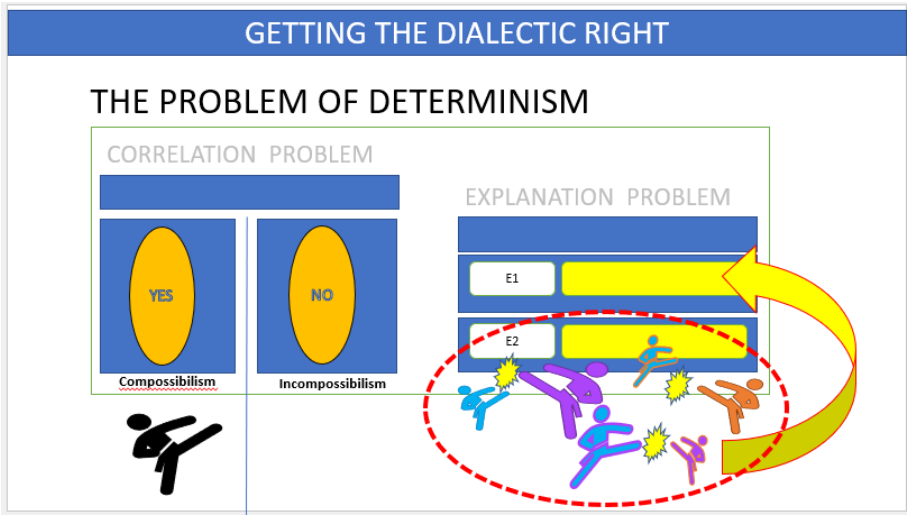
The Paradox of (In)determinism








THE PROBLEM OF DETERMINISM
CORRELATION PROBLEM
EXPLANATION PROBLEM
YES
NO
E1
??
E2
Causal Luck Incompossibilism
∴ Constitutive Luck Incompossibilism
∴ Circumstantial Luck Incompossibilism
∴ Hybrid Incompossibilisms
COMPOSSIBILISM : INCOMPOSSIBILISM


GETTING THE DIALECTIC RIGHT
THE PROBLEM OF DETERMINISM
CORRELATION PROBLEM
EXPLANATION PROBLEM
YES
NO
E1
E2
Compossibilism
Incompossibilism


THIS IS KAREN...
SHE'D LIKE TO SPEAK TO YOUR MANAGER ABOUT THE CORRELATION/RELEVANCE DISTINCTION
Nobody likes a Karen, Karen.

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,

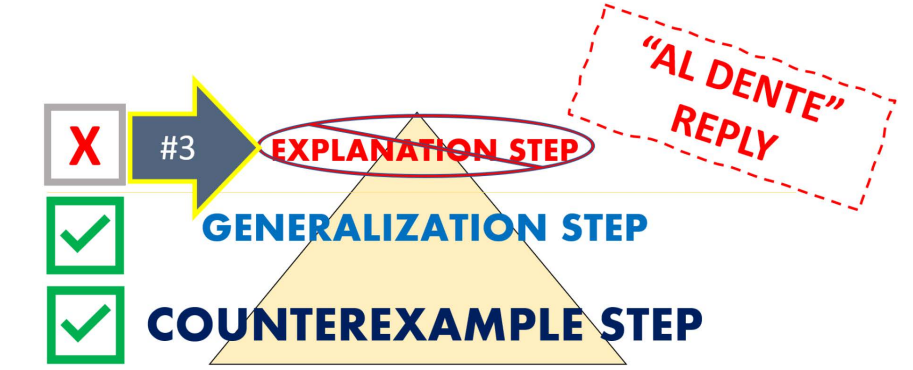
MY

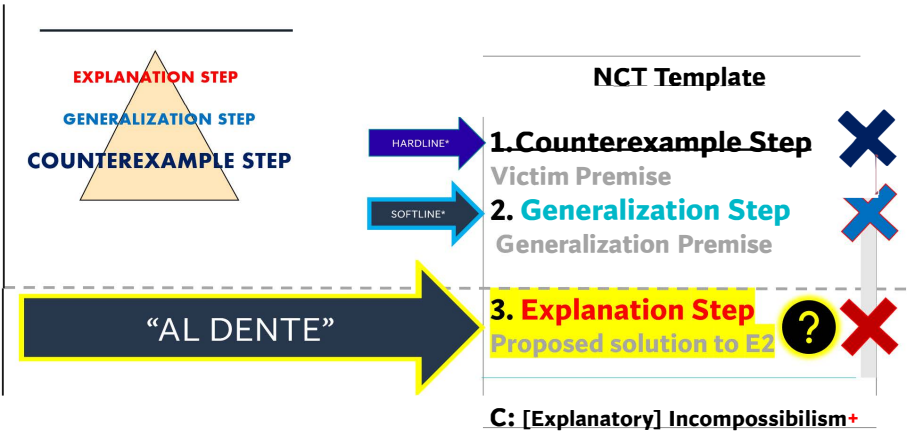
al dente reply to

OZA-2006 / OZA-2012 shows that:

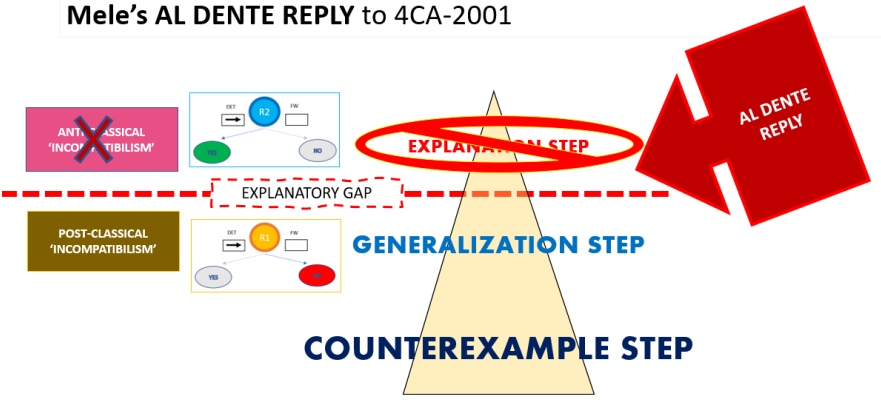
Mele's Original Zygote Argument for *incompatibilism* fails.

EXPLANTORY 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



FOR THE RECORD

ALFRED R. MELE

ASPECTS OF AGENCY 6

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]

9
Manipulation and Direct Arguments
JUSTIN A. CAPES
A COMPANION TO
FREE WILL

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

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

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

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 t
premise 1 nor 2 identifies deterministic laws as "menacing," i.e. fre
responsibility-undermining.[1] Yes, Ernie lives in a deterministic uni
premise 1 does not assert that it is *in virtue of* being subject to determi
that Ernie lacks free will. For all that is said in the premises of ZA, th
Ernie's universe is deterministic may or may not be relevant to his statu
agent. If anything, the emphasis in premise 1 of ZA is on the way t
zygote was *produced*. However, Mele's description of ZA makes it cle
should not read premise 1 as promoting a particular account of Ernie
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. Incompossibilism
2. So, incompossibilism 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 "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.