

Bartłomiej Andrzej Lenart*

Metaphysical Compatibilism and the Ontology of Trans-World Personhood: A Neo-Lewisian Argument for the Compatibility of Divine Foreknowledge (Determinism) and Metaphysical Free Will

<https://doi.org/10.1515/mp-2021-0041>

Published online July 13, 2022

Abstract: David Lewis' contemplations regarding divine foreknowledge and free will, along with some of his other more substantial work on modal realism and his counterpart theory can serve as a springboard to a novel solution to the foreknowledge and metaphysical freedom puzzle, namely a proposal that genuine metaphysical freedom is compatible with determinism, which is quite different from the usual compatibilist focus on the compatibility between determinism and moral responsibility. This paper argues that while Lewis opens the doors to such a possibility, in order to fully elucidate a genuinely *metaphysical compatibilist* account, Lewis' own counterpart theory must be abandoned in favour of an account of trans-world identity that is theoretically framed by a modified version of Robert Nozick's closest continuer theory.

Keywords: compatibilism, metaphysical compatibilism, free will, determinism, divine foreknowledge, trans-world identity, multiverse, modal realism, counterpart theory, closest continuer theory

1 Introduction

Augustine pondered: how can it be that God has foreknowledge of all future events, and yet that we do not sin by necessity" (Augustine 391/2002, p. 260)? The problem of foreknowledge and metaphysical freedom is not exclusively the domain of theology, but is rather, more generally, a question regarding the possibility of metaphysical freedom within the context of a deterministic universe. One

*Corresponding author: Bartłomiej Andrzej Lenart, University of Calgary, Calgary, AB, Canada, E-mail: bartlomiej.lenart@ucalgary.ca

solution to this problem is to presuppose a *presentist* ontology, which states that neither past nor future events actually exist, and thus that statements can only be true of the present. This line of argument entails that neither divine foreknowledge nor determinism fix the truth values of statements in the future, meaning that no proposition with regard to future contingencies carry any truth values. See Johnson (2009) for this type of argument. Presentism, however, is difficult to defend against the objection that the ontological claims of the proposal are either false when the view is understood as making the claim that only present things have, do, or ever will exist, and merely trivial when the view is taken to simply state that only present things exist now. For a discussion of this sort of argument, see Ludlow (2004) and Crisp (2004).

The most commonly adopted solution to the above problem, especially when stated in contemporary terms, is compatibilism (or soft determinism), which makes the claim that free will is compatible with determinism as long as agent actions are caused by agent desires. Unfortunately, compatibilism, in this form, fails to capture the proper sense and meaning of the term ‘free will,’ especially considering that, if determinism is true, then agent desires are also determined. Thus, compatibilists mostly focus on the compatibility between determinism and moral responsibility. Insofar as this is the case, this type of compatibilist proposal ought better to be termed *moral compatibilism*, in order to distinguish it from another, stronger, compatibilist notion, which I shall term *metaphysical compatibilism* (following van Inwagen’s usage of the term *metaphysical freedom*¹), that makes the claim that genuine metaphysical freedom is possible even in the context of divine foreknowledge (which entails a deterministic worldview). The above disambiguation is a means of distinguishing two general, but sometimes conflated, meanings of compatibilism: (1) the strong, metaphysical usage, which states that *metaphysical freedom* is compatible with determinism, and (2) the weaker, moral usage, which focuses on the compatibility of determinism with moral accountability.

A further note on terminology usage in this paper: when referring to trans-world identity in this paper, I am focusing on what some have termed trans-world individuals; trans-world identity can be taken to more generally refer to the notion of something existing in more than one world, which is different from the idea of something existing in more than one world in virtue of having different parts in different worlds. When using the term trans-world identity, I am referring to this latter concept more specifically.

¹ I am following Peter van Inwagen (1998) in the usage of the term *metaphysical freedom* to refer to a robust libertarian conception of free will.

While it may be true that David Lewis (1976a) paper “Are we free to break the laws?” does not actually explain how to be a compatibilist, but rather only suggests that there is no good reason for incompatibilism, I think it is possible to construct an account of compatibilism on Lewis’ behalf that is reasonably consistent with what he argues in the above mentioned paper (as well as elsewhere). Moreover, not only is Lewis a compatibilist (see Lewis 1999, p. 112), but Lewis’ works offer enough material to construct a worthwhile and robust compatibilist proposal, which can fairly be termed *metaphysical compatibilism*. While he does not fully explain his own approach to compatibilism, a charitable reading of his various works lends itself to a reconstruction of an intriguing compatibilist approach.

To this end, this paper engages in a reconstruction of a possible Lewisian account of compatibilism that offers a strong version of compatibilism all too often ignored in the free will literature. This paper is not merely reporting on Lewis’ own philosophical stance on compatibilism, but rather serves as an exploration of a possible Lewisian compatibilist account, one that is both plausible as well as charitable within the context of Lewis’ greater body of work, but also one that diverges from Lewis’ own position. My proposal also serves to answer the sort of question with which Lewis begins his paper “Evil for freedom’s sake,” namely whether divine foreknowledge (which assumes a predetermined state of events) can be compatible with metaphysical freedom. I believe that there is a way to construct such an account by following Lewis’ own train of thought some considerable way before diverging somewhat from Lewis’ own commitments and adopting the notion of trans-world identity in favour of Lewis’ own counterpart theory. I argue that all this can be accomplished within a slightly modified framework of Robert Nozick’s (1981) Closest Continuer Theory.

2 Argument Overview

This paper argues that the conjunction of David Lewis’ modal realism and his counterpart theory, as applied to personal identity across worlds, prohibits Lewis from interpreting the counterfactual statement ‘*could have done otherwise*’ in what this paper refers to as a *metaphysical compatibilist* manner because of Lewis’ own rejection of the notion of trans-world identity. Nevertheless, I argue that Lewis has set us on the right path toward a stronger compatibilist stance, even if his own theoretical commitments never allowed him to take the arguments he set forward to their logical conclusion.

The limitations of the Lewisian approach reside in the fact that while Lewis proposes that small miracles occur within possible worlds, neither his modal realism nor his counterpart-theoretic conception of personal identity allow or require such small miracles to comprehensibly describe possibilities within worlds or similarities across possible worlds. Although this paper makes heavy reference to Lewis' modal realism and counterpart theory in an attempt to unpack his philosophical stance on the metaphysical problem of free will, the purpose of this paper is to argue for a novel compatibilist thesis, which while inspired by the Lewisian approach, was never endorsed by Lewis himself and cannot be fully realized without some major departures from Lewisian metaphysics.

I argue that while Lewis' counterpart theory correctly tracks the modal properties of objects, his modal realism does not lend itself to solving the problem of metaphysical freedom in deterministic universes. Building on Saul Kripke's (1980) objection to Lewis' counterpart theory, I propose that Lewis' own notion of *perdurance* as applied to personal identity, individuates counterpart persons in the same way Kripke argues Lewis' view individuates other counterpart objects; this is because the tracking of diachronic personal identity across Lewisian worlds, in virtue of the fact that Lewisian persons reside in spatiotemporally isolated worlds, is not a matter of *branching* (as would be true for trans-world identity), but is rather a case of *divergence* of fully individuated spacetime entities merely sharing similarity relations between counterfactually indistinguishable spatiotemporal parts.

I further argue that this observation entails that just as the individuated possible worlds are isolated and deterministic in nature (in virtue of their adherence to their respective sets of laws), so persons are individuated in just this same way (in virtue of their relationships to diverging person-branches as described by the notion of perdurance). Therefore, without an appeal to some sort of trans-world identity (which Lewis explicitly denies), Lewis' account of persons, coupled with his counterpart theory (understood in light of the Kripkean objection), entails that persons cannot be metaphysically free actors in the kind of ontological framework Lewis postulates. This, however, does not preclude us, multiverse dwelling philosophers, from utilizing Lewis' insightful contemplations as a springboard toward a more robust conception of compatibilist freedom.

3 Mapping the Theoretical Landscape of the Free Will Debate

The philosophical problem of free will may be conceptually mapped in terms of divergent approaches to unpacking the statement, '*could have done otherwise.*' To

the libertarian, the statement ‘*could have done otherwise*’ carries the metaphysical implication that the agent can genuinely choose between either action X or action Y regardless of the environmental conditions, mental states, and laws of nature at play. This genuine ability to choose between actual opportunities is sometimes referred to as *metaphysical free will* or *metaphysical freedom*. For examples of libertarian arguments, see Chisholm, 1964/2004, Kane (1999/2004, 2000), and van Inwagen (1983, 1998).

Conversely, determinists understand the statement ‘*could have done otherwise*’ in counterfactual terms. To the determinist, the agent could only have done X given certain environmental conditions, mental states, and laws of nature – none of which are freely chosen by the agent. This same agent can only do Y under different conditions, the influence of different mental states, or given divergent laws of nature. In other words, the agent has no choice between X or Y due to the influence of these external factors and therefore cannot be held morally responsible for her choice. The environmental conditions, mental states, and laws of nature dictate which action the agent performs. For examples of determinist positions (also known as hard-determinist stances), see Holbach (1770/2005), Pereboom (2001, 2005), and Wegner (2003).

The libertarian thesis is incompatible with causal determinism because it disregards the binding causal forces of both external and internal causal mechanisms by postulating an independent metaphysical process of agency, which is causally free from the influence of such things as environmental conditions, inner mental or psychological mechanisms, or the laws of nature. Compatibilists often try to position themselves somewhere in between the two conceptual camps of libertarianism and determinism.

Hobbsian compatibilism essentially states that an agent is free when she is doing what she desires or intends to do. Hobbes writes:

A FREE-MAN, is ‘he, that in those things, which by his strength and wit he is able to do, is not hindered to doe what he has a will to . . . from the use of the word Freewill, no liberty can be inferred to the will, desire, or inclination, but the liberty of the man; which consisteth in this, that he finds no stop, in doing what he has the will, desire, or inclination to doe. (Hobbes 1651/2018, p. 200)

Compatibilists who follow the Hobbsian line of argument, therefore, contend that when agent desires and inclinations happen to align with agent actions, the agent is acting freely even if determinism is the case. There is, nevertheless, a real sense in which, on such a view, the agent could not have done otherwise than she actually did, meaning that there is no sense in which an agent could have changed her mind or done something other than what she ended up doing. This kind of compatibilist freedom, then, is what Wegner (2003) refers to as illusory

freedom. What compatibilists like Saul Smilansky (2000) argue is that this type of illusory freedom is enough to justify our common sense judgments regarding accountability and moral responsibility.

Another common classical compatibilist approach is to analyze ‘*could have done otherwise*’ in conditional terms (see Ayer 1954; Hobart 1934, and Hume 1784/1975). The idea is that ‘*could have done otherwise*’ merely means that while an agent did X instead of Y, she did X because she wanted to do X. It is worth noting the classical compatibilist language here, which while it is not compatible with genuinely having the ability to do otherwise, nevertheless attempts to capture a compatibilist notion of free willing. A conditional analysis of ‘*could have done otherwise*,’ then, states that the agent could have done Y had she wanted to do Y instead of X, where wanting to X or Y is nevertheless still determined by the agent’s psychology and environment, and thus is not up to the agent. What such conditional analyses of ‘*could have done otherwise*’ amount to is simply stating that if things had happened differently and the laws of nature and psychological states of the agent had been different, the agent would have done otherwise. This, however, is a far cry from the kind of genuine ability to do otherwise captured by libertarian intuitions, and thus merely amounts to counterfactually stating that if different conditions existed, the agent would have been determined to act differently or to make a different, albeit nevertheless still very determined, choice. Such approaches, unfortunately boil down to mere semantic tricks and fall far short of pointing to a meaningful and robust sense of metaphysical freedom within a deterministic universe.

Compatibilism, generally speaking, in virtue of being concerned with the ultimate causal forces of agency, is, by extension, primarily concerned with the question of ultimate moral accountability for actions that are, under a determinist worldview, not up to the agent. “Because free will is typically taken to be a necessary condition of moral responsibility, compatibilism is sometimes expressed as a thesis about the compatibility between moral responsibility and determinism” (McKenna and Coates 2021, introduction).

Two major contributions to the free will question drastically changed the course of compatibilist thinking. The first was Peter Strawson’s argument published in the early 1960s, which pointed out the significance of interpersonal relations to the practice of holding people accountable for their actions. The second was an attack on the Principle of Alternative Possibilities by Harry Frankfurt published in the late 1960s and early 1970s.

Peter Strawson (1963/2004) classic argument points out that moral accountability is compatible with determinism due to the attributions of reactive attitudes that ground responsibility in the agent’s actions regardless of whether or not the

actions themselves are metaphysically free. The reason for this is that such reactive moral attitudes are the grounding for all of our social and interpersonal relationships, and thus cannot reasonably be abandoned. See Lenart (2007) for a discussion of Strawson's notion of reactive attitudes within the context of free will and computational theories of mind.

Harry Frankfurt (1969, 1971) argues that from the perspective of a compatibilist, '*could have done otherwise*' is merely a means of attributing the responsibility and moral consequence of choice to the agent. According to Frankfurt, the agent is morally responsible for her action if she still *would* have done X even if she genuinely *could* have done Y – this still holds, however, even if she genuinely *could not* have done Y. The agent remains morally responsible for her action by virtue of her second order desire to do X (which is the desire to have the desire for X). The agent is not free from her moral responsibility simply because Y was not genuinely available to her. "The point of Frankfurt's paper was to drive a wedge between responsibility and alternate possibilities, and to thereby drive a wedge between responsibility and libertarian freedom" (Zagrzebski 2021, Section 2.5).

Compatibilist arguments, which propose the stronger compatibilist thesis that metaphysical freedom is compatible with determinism, ought to be distinguished from the moral compatibilist positions (like those of Frankfurt and Strawson), which state that the lack of metaphysical freedom within a deterministic universe is nevertheless compatible with the social practice of holding agents morally accountable for their actions.

The Lewisian compatibilist stance, as outlined in his 1981 paper "Are we free to break the laws?," attacks the first premise of what McKenna and Coates (2021) term the Consequence Argument² by asserting that rather than having the power to violate the laws of nature, the compatibilist need only postulate that if the agent were to act otherwise than she did, some law of nature, which actually obtained would have had not to obtain. Lewis can track this response to the Consequence Argument via his modal realist analysis of counterparts in other worlds. Unfortunately, as I will argue, while this approach is promising, counterpart theory can only account for a weaker version of the compatibilist claim, one which, at best, can only have significance with regard to the question of moral accountability, but does not, in fact, make any stronger metaphysical claims. I will also argue that if counterpart theory is replaced with the notion of trans-world identity, this neo-Lewisian approach gains some much needed metaphysical traction.

2 1. No one has power over the facts of the past and the laws of nature. 2. No one has power over the fact that the facts of the past and the laws of nature entail every fact of the future (i.e., determinism is true). 3. Therefore, no one has power over the facts of the future (McKenna and Coates 2021, Section 4.1).

Lewis himself understood that compatibilist freedom in the sense that both Frankfurt and Strawson tackled the term has very little relation to metaphysical freedom. Lewis writes: “I’ll speak of ‘compatibilist freedom’ and ‘incompatibilist freedom’. But I don’t ask you to presuppose that these are two varieties of freedom. According to incompatibilism, compatibilist freedom is no more freedom than counterfeit money is money” (Lewis 1999, p. 110). As I will explain in much greater detail below, Lewis’ argument for small miracles, as outlined in his paper “Are we free to break the laws?,” can be read as an extension of the thought experiment in his (Lewis 1976b) “The paradoxes of time travel” where, in an attempt to resolve the grandfather paradox,³ Lewis proposes the solution that branching would need to occur to make proper sense of the grandson’s paradoxical capacity and conjoined lack of ability to kill his own grandfather.

Since, in the case of Tim and his grandfather, Tim’s timeline branches into two possibilities, we can assume that this branching (or divergence) of timelines can be tracked by a divergence of worlds. Given that Lewisian worlds are causally isolated, the small miracles can serve as an account of how to make sense of the divergence of Tim’s worlds in a way that speaks to Tim’s ability to make the free choice to kill his grandfather, even if, for all intents and purposes, the paradox of time travel can only be avoided if Tim’s choice results in world branching. While Lewis is not arguing for Tim’s libertarian freedom, he is, I argue, setting the stage for a brand of compatibilism that does not suffer from the sort of counterfeit freedom Lewis rejects in his 1999 paper. Although Lewis himself does not take this idea to its logical conclusion due to his commitment to counterpart theory, the stage is set for such an account to emerge.

4 Trans-World Identity and Perduring Persons

“The notion of transworld identity — ‘identity across possible worlds’ — is the notion that the same object exists in more than one possible world (with the actual world treated as one of the possible worlds)” (Mackie 2017, Introduction). When applied to persons, the view identifies counterparts at different possible worlds with one another. That is, copies of an individual are VIEWED as parts of a much larger trans-world self. For more on trans-world identity see: Lenart (2019), Lewis (2004), McDaniel (2004), Papineau (2004), and Yagisawa (2010).

³ The grandfather paradox states that there is a sense in which a time travelling grandson who travels back to a time before his own birth both can and cannot kill his own grandfather.

The problem with trans-world identity, according to Lewis (1986), is that trans-world individuals lead very peculiar double lives. That is, a trans-world individual somehow manages to have contradictory modal properties. For instance, Humphrey at world W_1 might win the presidency while Humphrey at W_2 does not win. Thus, the thesis of trans-world identity states “that the *whole* of Humphrey is part of different worlds, with different properties at different ones” (Lewis 1986, p. 200). Lewis’ discomfort with the notion that a single individual can be in possession of contradictory properties prevents Lewis from entertaining and adopting the trans-world identity proposal.

Thus, while Lewis does postulate that worlds can share parts, he does not share in the trans-world sentiment that persons can.⁴ He writes:

Given the unrestricted mereology I favour, sharing of parts is altogether commonplace. Indeed, any part of any world is part of countless mereological sums that extend beyond that world. But what I do find problematic . . . is the way the common part of two worlds is supposed to have different properties in one world and in the other. (Lewis 1986, p. 199)

Later in the same chapter, Lewis states: “I shall argue that indeed there are things that enjoy trans-world identity in this sense. But then I shall argue that we ourselves, and other things that we ordinarily name, or classify under predicates, or quantify over, are not among them” (Lewis 1986, p. 210).

Counterpart theory, on the other hand, Lewis argues, allows for the differentiation of each counterpart’s intrinsic properties because counterparts are not exact intrinsic duplicates. Individuals, on Lewis’ view, *perdure* through time (they are made of temporal parts), but not through worlds because the unification of the parts proves more problematic for modality than for time – see Lewis (1983) for a full account of perduring persons. This solves the problem of temporary intrinsics since even though temporal parts differ from one another (i.e. person S is young, has hair, and walks up straight at time t_1 , but is old, bald, and hunched over at time t_2), unified, trackable, and individuated persons are made up of these differing temporal parts; that is, a person is the aggregate of these parts and each part is merely a person stage and not the entire person. According to Lewis, the reason why individuals perdure through time, but not through worlds is that: first, there can be no trans-world causation that unites counterparts (worlds, on Lewis’ view, are spatiotemporally, and thus causally, isolated); second, the unification of trans-world individuals should be a matter of direct similarity between the (trans-world) person stages (but this is impossible because change across worlds is not really limited to any particular chain of person stages); and third, whereas

⁴ While Lewis does argue that persons can share temporal parts within a given world, he rejects the trans-world identity proposal that a person can share different parts in different worlds.

single-world fission and fusion puzzles are rarities entertained by science fiction writers and philosophers (and thus are not pathological), “modality is different: pathology is everywhere” (Lewis 1986, p. 219) because worlds diverge and converge in *all* possible ways. Finally, “[t]he supposed trans-world person, no matter how well unified by counterpart relations, is not the sort of integrated self that is capable of self-interest” (Lewis 1986, p. 219) because whereas a person perduring through time is still capable of collective self-interest, the trans-world self is faced with too many contradictory desires and intentions and thus “[m]y this-worldly self has *no* tendency to make the purposes of its other-worldly counterparts its own” (Lewis 1986, p. 219).

A noteworthy distinction to keep in mind is between branching and divergence: Lewis’ worlds do not branch, but rather merely diverge. Branching entails the coming to be of two different futures from one initial spatiotemporal segment. Such futures differ both numerically and qualitatively. Fission and overlap cases are therefore problematic for both worlds and individuals. “In divergence, on the other hand, there is no overlap. Two worlds have two duplicate initial segments, not one that they share in common. I, and the world I am part of, have only one future” (Lewis 1986, p. 206). Lewis’ treatment of single-world fission (e.g. in duplication cases) follows a similar line of thought – one which, I think, only fortifies the notion that counterparts in other worlds are separate and unique even if they happen to be eerily similar in many ways. In “Survival and identity,” Lewis considers both fission and fusion cases where non-identical continuants C_1 and C_2 both share person-stages either before fission or after fusion. Lewis argues that C_1 and C_2 are individuated all along (even during the period of identity) and that they only happen to share some person-stages. Thus, Lewisian individuals (like Lewisian worlds) do not branch, but rather diverge. However, Lewis points out that it may be practically advantageous to refrain from individuating C_1 from C_2 while they share stages since the two individuals have the same properties and thus, for all intents and purposes, can be treated as identical; nevertheless, this identity relation is assumed only for practical purposes, rather than as a metaphysical or ontological means of individuating and tracking personhood.

5 Modal Realism and Small Miracles

What Lewis means by determinism is that “the prevailing laws of nature are such that there do not exist any two possible worlds which are exactly alike up to some time, which differ thereafter, and in which those laws are never violated” (Lewis 1973, p. 559). Thus, in “Are we free to break the laws,” he assumes that the ability to do otherwise requires small miracles. However, Lewis’ modal realism

commits him to an analysis of counterfactuals that is framed in terms of possible worlds. Consequently, ‘*could have done otherwise*’ for Humphrey in the actual world $W@$, means that Humphrey’s counterpart S_1 in a very similar possible world W_1 , did otherwise. So, even though Humphrey did not raise his arm or win the presidency in $W@$, Lewis maintains that he *could* have. But since Lewis rejects trans-world identity, he must interpret this *could* in a counterpart theoretical manner, which, in turn, leads him to the postulation of the necessity of small miracles.

The view is as follows: S_0 (Humphrey) at W_0 (a world with a given history H and laws of nature L) could have done x if S has a counterpart S_1 who did x at a possible world W_1 , which is qualitatively (even though not numerically) identical to W_0 (that is, W_0 and W_1 have the same history H and laws L , but are numerically and causally separate). The act of x -ing at W_1 causes W_0 and W_1 to diverge. And so, Lewis suggests, since H & L are true of both worlds, the divergence must be caused by a breakage of laws. That is not completely accurate, however, since Lewis uses the term “miracle” to express a relation between the worlds where the laws L in both worlds are, in fact, *almost* identical, meaning that while H may be the same in both W_0 and W_1 , L is not identical in both worlds; rather L underlies the evolution of events in W_0 while some other, very similar “almost-laws” L^* govern the unfolding events in W_1 . Lewis explains: “[a] miracle at w_1 , relative to w_0 , is a violation at w_1 , of the laws of w_0 , which are at best the almost-laws of w_1 . The laws of w_1 itself, if such there be, do not enter into it” (Lewis 1979, p. 469). The above explanation clearly illustrates Lewis’ commitment to the causal separation of worlds. Moreover, it also hints at Lewis’ ontological preference toward understanding the unfolding histories in all possible (and causally separate) worlds in deterministic terms since, according to the above cited passage, the unfolding of events within W_0 and W_1 are necessarily linked to the respective sets of laws in each world, where the mutability of laws (i.e. the occurrence of small miracles) does not actually take place within each respective world, but rather only exists as a comparison between W_0 and W_1 . This, to my mind, does not constitute a breakage of laws or even a small miracle within either world (W_0 and W_1), and is consistent with both W_0 and W_1 being governed by deterministic laws (albeit L and L^* respectively).

To turn to one of Lewis’ own examples: Humphrey, who *could* have won the presidency (or *could* have raised his arm), but did not, satisfies the formula ‘ $\diamond x$ wins,’ but not the formula ‘ x wins’ since Humphrey does not actually win the presidency in $W@$. This means that there is some world W such that at that world, Humphrey satisfies ‘ x wins.’ Lewis, however, rejects the notion that Humphrey is part of W as well as a part of $W@$ for reasons outlined above. Thus, Lewis is compelled to say that Humphrey need not be part of a world to satisfy a formula

there. Therefore, according to Lewis, “there is a world where somehow he satisfies ‘x wins’ *in absentia*” (Lewis 1986, p. 9). On Lewis’ counterpart-theoretic account, satisfaction *in absentia* is vicarious satisfaction. That is, “Humphrey satisfies ‘x wins’ vicariously at any world where he has a winning counterpart” (Lewis 1986, p. 10).

This is somewhat puzzling, however, since, following Kripke’s (1980) objection to Lewis and Lewis’ own statements regarding miracles (Lewis 1979, p. 469), it is unclear how Humphrey can satisfy *x in absentia* (or vicariously) if Humphrey’s counterpart is a distinct individual (as has been argued above). Kripke writes:

if we say ‘Humphrey might have won the election (if only he had done such-and-such),’ we are not talking about something that might have happened to *Humphrey*, but to someone else, a ‘counterpart.’ Probably, however, Humphrey could not care less whether someone *else*, no matter how much resembling him, would have been victorious in another possible world. (Kripke, 1980, p. 45)

In order to further buttress Kripke’s objection, let us examine a different, but somewhat related, case of single-world counterparts. In a single world (W) where Humphrey (T_1) has an identical twin (T_2), both T_1 and T_2 are separate individuals (as, according to Lewis, is the case for Humphrey and his counterparts in other worlds). If both T_1 and T_2 run for the presidency in W and T_2 wins, but T_1 does not, it is absurd to say that T_1 satisfies ‘x wins’ either *in absentia* or vicariously. Although T_1 might be happy for his brother and, in fact, live and enjoy the presidency, in some sense, vicariously through his brother’s success, it does not make very much sense to say that T_1 counterfactually satisfies the formula ‘x wins’ merely in virtue of the fact that he has a counterpart (his twin) who actually wins.

What, then, makes counterparts in other worlds so special (especially if they are even more removed from Humphrey in $W@$ than T_1 is from T_2 in W)? Perhaps the other worlds themselves are special somehow and thus, Humphrey satisfies ‘x wins’ *in absentia* in virtue of the very existence of other worlds? Unfortunately, I do not see how this has any bearing unless we identify Humphrey with his counterpart via some sort of trans-world identity relation, but Lewis’ rejection of trans-world identity bars him from a more intimate identification between counterparts.

It could be argued that since Humphrey and his counterpart share the same history, they are, in fact, much more alike than are T_1 and T_2 . The point, however, is that Humphrey and his counterpart (just like T_1 and T_2) are distinct individuals regardless of whether or not they share the same history because the sharing of history (both in single-world fission and fusion cases as well as in cases of divergence and convergence of worlds), in fact, reduces to merely sharing person-stages in the single-world case and it reduces to merely a relation of similarity

across possible (or counterfactual) worlds in the many-worlds case. After all, on Lewis' account, counterparts in other worlds are not only distinct individuals, but they inhabit spatiotemporally isolated worlds.

Humphrey's satisfying 'x wins' *in absentia* suggests that Humphrey satisfies both 'x wins' and 'x does not win' (even if each is satisfied in a separate world). Thus, all the while keeping in mind that Lewis' worlds are all actual (concrete), because he is a modal realist, either we are entering the realm of trans-world individuals where Humphrey satisfies two contradictory propositions (which Lewis cannot accept) or Humphrey, in the actual world (where he does not win), should simply be accused of wishful thinking. This is because, in both the single-world as well as the other-world cases, respectively, $T_1 \neq T_2$ and $S_1 \neq S_2$ (unless, as already mentioned, in the other-world case, we succumb to the temptation of trans-world identity). It seems to me that there should be more than just the relation of similarity unifying counterparts for an appropriate analysis of 'could have done otherwise.'

A more technical way of formulating Kripke's (1980) worry goes as follows: a possible object, for Lewis, does not exist in the actual world ($W@$), but it does exist in a possible world (W_1). Note that '@' is world-relative in that it works like an indexical (that is, every world is *actual* from that world's point of view and every world is *possible* from another world's point of view). And so the formula $\diamond(\exists x)Sx$ (where S is the predicate "is my sister") in $W@$ (where there is no x that satisfies S) cannot be satisfied, on Lewis' view, if we restrict the quantification to $W@$. This is why Lewis quantifies over all the worlds and thus $\diamond(\exists x)Sx$ in $W@$ just means $(\exists x)Sx$ in W_1 since there is no 'a' such that Sa in $W@$, but there does exist an 'a' such that Sa in W_1 . However, if Lewis rejects trans-world identity in favour of counterpart theory, how can we sensibly talk about *my* possible sister on the modal realist view if, in order to track the possible object (Sa), we are forced to quantify over other worlds where I do not exist, but merely a counterpart of *me* exists? Doesn't this suggest that $\diamond(\exists x)Sx$ in $W@$ and $(\exists x)Sx$ in W_1 track different objects (since counterparts are separate and individuated)? That is, $\diamond(\exists x)Sx$ refers to *my* possible sister, but 'my sister' at W_1 is the sister of my counterpart (and thus, the 'my' in 'my sister' is an utterance made by the counterpart in W_1 and *not* by *me* in $W@$). A similar puzzle arises for a counterpart-theoretic analysis of 'could have done otherwise.'⁵ I turn to this problem in the following section.

⁵ My intuition is that Linsky and Zalta's (1994) actualist semantics fares better (at least their account appears to be more intuitive than Lewis') because rather than quantifying over many worlds and many different and separate counterparts, they quantify over the actual world ($W@$) on their view.

Lewis replies to Kripke's objection as follows:

Counterpart theory does say . . . that someone else – the victorious counterpart – enters into the story of how it is that another world represents Humphrey as winning, and thereby enters into the story of how it is that Humphrey might have won . . . What matters is that someone else . . . should not crowd out Humphrey himself . . . Thanks to the victorious counterpart, Humphrey himself has the requisite modal property: we can truly say that *he* might have won. There is no need to deny that the victorious counterpart also makes true a second statement describing the very same possibility: we can truly say that Humphrey-like counterpart might have won. (Lewis 1986, p. 196)

While this may be true for Humphrey's modal properties across worlds, Kripke's objection still stands with regard to Humphrey's mereological properties; that is, while Lewis' modal realism accurately tracks Humphrey's modal properties, it cannot be applied to Humphrey's mereological properties since, by Lewis' own argument, Humphrey is not the kind of individual who can share different parts with his counterparts across different worlds. The mereological properties of persons⁶ are of interest to the question of free will. Since Lewis argues for local miracle compatibilism on the grounds of the divergence of identical worlds, it is reasonable to conclude that an agent's modal properties track an agent's possible actions across worlds. However, given that an agent's metaphysical properties are world-specific (and, unlike moral properties, cannot be tracked across worlds), an agent's actions cannot similarly be tracked across worlds, even though her possible or probable actions may be tracked modally across worlds in the way Lewis suggests.

6 Why Lewis Cannot Be a Metaphysical Compatibilist

While, as already explained above, my starting point is Lewisian in genesis, Lewis himself is not a metaphysical compatibilist due to at least two major theoretical commitments. This section outlines why Lewis does not subscribe to the metaphysical compatibilist stance, and the following section proposes to diverge from the Lewisian origin as it offers to outline a more detailed account of my proposal.

Lewis' first theoretical commitment entails that each small miracle is only a miracle relative to a comparison between causally separate worlds, which does not, in fact, amount to any ontological anomalies that would appropriately be dubbed "small miracles" in the way required for the postulation of a

⁶ These persons are agents in their respective worlds.

robust strong compatibilist thesis. Therefore, the nomological differences between worlds merely amount to differences between isolated worlds with undisrupted nomological histories. Divergence between “almost-identical” worlds can at most be the result of either some hidden nomological variables or probabilistic laws; neither option, however, opens the door to metaphysical freedom since the first still results in a deterministic worldview, and the latter entails chance – neither of which is compatible with metaphysical freedom; for discussions regarding metaphysical freedom and the problem of indeterminism, probability, and chance see Kane (1999, 2000), Lenart (2007), Nozick (1981), and Searle (2004).

Similarly, Lewis’ second theoretical commitment (that of perdurance) also amounts to a view of persons where each counterpart is a wholly separate (isolated) individual, and thus has only one future and is, as it were, stuck with whatever action he or she undertakes (regardless of what his or her numerous counterparts may or may not do). Moreover, assuming causal determinism (at a particular world), the agent’s actions are entirely dependent on the agent’s past and the laws of the world the agent happens to inhabit because these laws and the agent’s history (as well as the history of the world) are features of that particular world. Very similar, but divergent worlds have laws and histories that are *almost identical* to the laws and history of $W@$, but are, in fact, neither numerically nor qualitatively *identical*. Possible worlds only ever diverge on Lewis’ view, which entails that they were different all along. Thus, it would seem that no small miracle is necessary to explain the divergence since while divergent worlds are very *similar*, they were never *identical*. And again, the same can be said for person (agent) counterparts in nearby worlds; they are similar, but not identical; they are different, separate, and, in an important sense, unique (and merely divergent, rather than branching) individuals inhabiting similar, but different, separate, and, in fact, isolated worlds.

Since Lewisian persons are already always fully in existence as perduring entities, there is a strong sense in which what it means to be *me* is that *I* have made (and will have made) the choices *I* have, and any divergence from the actions, *I*, as an agent, have or will have made, belong to someone different from me, who merely happened to have shared some temporal parts with me. This, however, also suggests that Lewisian persons are pre-determined to have the spatiotemporal shape that they end up having, meaning that it is not unreasonable to read such an account of agency and personhood in deterministic terms. Furthermore, because Lewisian persons also inhabit a modal realist ontology, they must be understood in Lewis’ counterpart theoretical terms. As already argued above, because possible worlds are causally isolated, and since Lewisian persons are perduring entities, all possible actions of such entities living in such worlds must also be causally isolated, meaning that their entire pasts and futures

are already fully defined regardless of the branching that one would expect to occur in virtue of personal agency; this is because every instance of branching is actually more accurately described as an instance of divergence of worlds. Given such ontology and account of persons, Lewis cannot, therefore, reasonably be read in metaphysical compatibilist terms; it seems much more plausible that Lewisian persons reside in hard-determinist worlds, which would make Lewis a compatibilist of some other (perhaps Frankfurtian or Strawsonian) sort.

Persons residing in Lewis' modal realist multiverse perdure through causally distinct, isolated, and divergent (rather than branching) worlds. While worlds may share uncanny similarities at various times, the counterparts residing in each isolated world are only related to one another via a similarity relation rather than an identity relation; while such similarities allow us to track possibilities, the actualization of such possibilities in nearby possible worlds (in virtue of the postulation of Lewis' modal realism) has no causal bearing on the actual evolution of the temporal parts that are necessarily constitutive of Lewis' perduring persons. Moreover, the ontological shape of Lewisian persons, in virtue of their perduring nature, is static and determined by their own spatiotemporal and modal realist coordinates, where each perduring individual is subject to the laws of their own isolated, and merely divergent, world.

7 A Case for Metaphysical Compatibilism and the Shape of Persons

In this section, I will outline the argument for metaphysical compatibilism by highlighting the necessary departure point from the Lewisian line of thinking about persons across worlds. This section also offers an overview of the ontology of trans-world persons in light of some of the theoretical commitments my account necessitates. In order to robustly explore the implications of metaphysical compatibilism, it may be useful to frame the problem at hand by returning to the afore alluded idea that there can nevertheless be room for genuine metaphysical freedom even in a universe where a powerful calculator or an all-knowing deity has perfect information and thus is capable of predicting with precise and perfect accuracy how events in the universe will unfold. I argue that in order to make sense of this *prima facie* paradoxical proposal, the metaphysical compatibilist can adopt the Lewisian approach, but ought to give up counterpart theory in favour of an account of trans-world identity proposed in this section.

In his *Philosophical explanations*, Robert Nozick (1981) offers a powerful account of diachronic personal identity, which, while at the time of writing focused on solving problems of fission and fusion of the kind Derek Parfit (1986)

later popularized, serves, I argue, as a fitting template for the sort of account of trans-world identity needed for a strong compatibilist view to work. Nozick introduces the notion of a subjective metric as follows:

In synthesizing itself the I does not merely include certain items; it also conceives itself as (under certain circumstances) incorporating specified future items or stages. The I's self-synthesis includes a self-conception which projects itself into the future (Nozick 1981, p. 105).

Nozick argues that persons have the power to form their own characters in virtue of their capacity to assign weights to choices and reasons. These weight attributions can be thought of as preference assignments, which transform some actions into character forming actions; Robert Kane (1999/2004) refers to these kinds of preference informed actions as self-forming actions. These weight attributions set precedents for future actions; they “set up a framework within which we make future decisions” (Nozick 1981, p. 297). Moreover, these self-forming or character-forming actions are self-subsuming, meanings that the decisions that are in fact self-forming are also made in accordance with the agent's preferences.

Nozick proposes that in cases of fission, an agent's metric (which is the result of a lifetime of self-forming choices) is used to determine which of the two resulting individuals is the closest continuer of the pre-fission person. Where the metrics align closely enough, survival and continuation of identity can be established. In other words, X at t_1 is identical to Y rather than Z at t_2 if and only if Y is both (1) a continuer of X , and (2) that Y is a closer continuer to X than Z , where the closeness of continuation depends on the relative closeness of the subjective metrics (or self-conceptions) of X , Y , and Z .

In the case of multiverse branching, such fission can be understood to literally occur, meaning that trans-world individuals fission all the time. Insofar as these self-forming choices form the grounding of a person's self-conception, it is possible to track different trans-world persons through the web of branching decisions scattered across the probabilities and possibilities that constitute an agent's truly open future. These self-formed self-conceptions, then, are what determine the relative metrics utilized by trans-world agents to identify their closest continuers across multiverse branches. To put this in Lewisian language, such agents can be understood as perduring through the multiverse (or in modal realist terms, through possible worlds) by pruning the branches of their own counterparts, which do not meet the criteria of closeness (or metric resemblances) necessary for inter-branch identification.

On my approach to trans-world personhood, then, Humphrey in $W@$ actually does have a sister in a genuine sense in virtue of having a sister in W_1 , the reason being that Humphrey now perdures through multiple worlds. It is true that in

virtue of this trans-world perdurance, Humphrey can have contradictory properties. However, many of such contradictory properties (properties that would be out of character for Humphrey in $W@$ because they do not align with his self-conception or metric) are pruned by the self-forming mechanisms of the modified Closest Continuer model outlined above; these properties are pruned because they belong to metrics that cannot be tracked as closest continuers of Humphrey in $W@$. Nevertheless, those contradictory properties that do remain within the bounds of the trans-world person (properties that are trackable via Humphrey's subjective metric in $W@$ because they are not out of character for Humphrey in $W@$) are not actually outside of ordinary human experience; people often have contradictory desires, are capable of holding contradictory beliefs, and can act contrary to their everyday patterns of behaviour in ways that are not out of character (e.g. proponents of non-violence can participate in combat sports, vegetarians in usual contexts can eat animal products in specified contexts,⁷ individuals reliable within professional circumstances can be unreliable in personal contexts, etc., all without sacrificing the integrity of their overall identities). Persons, on this view, are somewhat fuzzy entities, as they spread through the plethora of possibilities and claim some branches while pruning others, but it is precisely because of this fuzziness that we can make genuine statements about what someone *could have done*.

From the standpoint of $W@$, an agent *could have done otherwise* in a genuine sense because she did do otherwise in W_1 (because her trans-world self encompasses some of these contradictory properties, although only those that are not pruned by her self-formed self-conception). Insofar as the trans-world agent is a perduring entity over all the worlds that are being tracked via the closest continuer relation, her trans-world history includes all of the genuine possibilities open to her (it is this feature of trans-world beings that makes them somewhat fuzzy). However, from the standpoint of the phenomenologically closed worlds (each world is so closed in virtue of the various multiverse person parts being only conscious of a portion of their larger selves), the trans-world agent parts have a distinct phenomenological experience of making a choice. The various trans-world parts are thus genuinely free to make choices within the bounds delimited by the trans-world whole which is being held together via the closest continuer relation obtaining between the various branches in virtue of self-subsuming subjective metrics, which bind the trans-world parts into a larger, albeit somewhat fuzzy, trans-world whole.

⁷ This was true for the great Polish mountaineer, Wojciech Kurtyka, who while being vegetarian in everyday life, did, out of necessity, consume meat products on his Himalayan expeditions.

This view entails the conclusion that while all of Humphrey (the trans-world whole that he truly is) is predetermined to branch in exactly the way he does (and thus the trans-world Humphrey has a distinct ontology or “shape” from a similar, but different perduring trans-world agent, Humphrey*), the various trans-world parts, each with their own unique standpoints and phenomenological histories, have nevertheless genuinely chosen their spatiotemporal coordinates along some of the larger, trans-world Humphrey’s branches. On this account, an all knowing deity or a super calculator equipped with perfect information can still predict the patterns of probabilities within worlds and thus the full trans-world history of all the choices made across the worlds bound by the closest continuer relation; this can happen while Humphrey *qua* trans-world part or localized self (localized via his particular phenomenological standpoint within a certain set of branches) is genuinely making choices, and is so doing freely in a strong metaphysical sense. Moreover, the localized trans-world part’s own phenomenology of choosing tracks the very moment branching occurs within the larger trans-world whole.

This, of course, is by no means a libertarian account of free will since the trans-world whole is not free to branch in just any way, but is predetermined to take on the ontological shape it does. On the other hand, this is not a deterministic account either since the trans-world parts themselves do choose between genuine possibilities and do make those choices freely since the branches resulting from each choice can be genuinely tracked by the self-forming mechanism of the closest continuer relation. This account is, in fact, a genuinely compatibilist account of free will, and not just in the weaker, moral sense, but in a strong, metaphysical sense since it satisfies both the predictability and metaphysical freedom criteria; a fully informed calculator or deity can, in principle, predict the shape of the trans-world whole, while genuinely free choosing nevertheless occurs on the level of the trans-world parts.

8 Conclusion

The appeal of the account on offer in this paper is that the proposed mechanism maintains that genuine choices exist and that agents truly choose between possible alternatives, without introducing or resorting to something ineffable (like an Aristotelian unmoved mover). Moreover, this account, while remaining physicalistic, does not collapse into a reductionistic and mechanistic materialism, and thus leaves open the notion that something genuinely self-conceiving and self-affirming is happening at the level of mental causation, which can account for free choices in an otherwise deterministic, physical universe.

This is not an account of metaphysical (or libertarian) freedom since I am limited by my self-conception in the types of choices and branches that are constitutive of me, but my self-conception itself need not be reducible to a materialistic mechanism and thus need not be subject to the laws of nature in the same way that my body and bodily mechanisms are subject to the laws of nature. Moreover, there are genuine choices in the multiverse and it is only the act of choosing that splits the multiverse into branches. Thus, while given all the information about the world and about my self-conception, an all-knowing deity or super calculator would still know what I (the trans-world whole) am going to do (therefore making the claim of determinism true), I am still choosing amid genuine possibilities and the phenomenology of this choice is tracking an actual choice (one that is not merely illusory). That is, I (the trans-world part) not only feel free, but am free to make a choice in the world, even if an all-knowing deity knows what that choice will be before I make it since it can calculate the pre-determined ontological shape of my trans-world whole.

For readers interested in strong compatibilism's moral efficacy, the account can be framed in practical terms as follows: while it may not be out of character for an agent S to commit a crime, like robbing a bank, S nevertheless has a real choice, one that results in the branching of the trans-world self, where in W_1 , S doesn't rob the bank, even if S does indeed rob the bank in W_2 . In W_1 , however, the trans-world part S1 is not tried and jailed for the robbery because S1's free choice resulted in the W_1 branch where S1 does not rob the bank. S2, however, once caught, is tried and jailed in W_2 ; this is morally and legally justified on the grounds that S2 made a free choice to rob the bank, which resulted in the branch W_2 . While S (the trans-world whole) is determined by her own subjective metric to track both S1 and S2 as her trans-world continuers, and so S is determined to branch into W_1 and W_2 , S's overall (pre-determined) trans-world shape is the product of free choices of her trans-world parts. Thus, although a super calculator can predict the overall shape of S (making determinism true), the choices that result in the branching itself are genuine choices. Moreover, S's overall shape emerges from S's self-chosen self-conception, which glues some of S's multiverse counterparts together and prunes others. For example, while it may be in S's character to possibly rob the bank (hence S branches into both S1 and S2), it may be entirely out of S's character to do so violently (which is why S3, who does so in W_3 , is not a trans-world part of S, but rather a part of some other, merely similar, yet not identical, trans-world whole).

Readers may note that I depart from Nozick's closest continuer theory here, but that is as it should be. Nozick argues that in cases where a closest continuer cannot be established during a fission (this can happen if both Y and Z are equally close to their predecessor X), the original does not survive. However, in the context

of trans-world identity, such a postulate does not make sense since trans-world wholes are the result of branching and thus, by definition, require that the identity relation is tracked across instances where both branches are equally the closest continuers of the pre-branch individual.

The metaphysical compatibilist account proposed in this paper tracks a compatibilism that is not merely of the weaker, moral form, but rather of a stronger type where, in addition to the phenomenology of choosing, there is also the physical manifestation of the choice in the form of the creation of multiverse branches, which bundle into trans-world individuals in accordance with self-conceived metrics of persons who are delimited by the bounds of some possible worlds, but whose individuality emerges from the vast expanse of all possible worlds.

Acknowledgment: I would like to extend a special thank you to Alicia Adlington for her thorough editing of an earlier (very different) draft of this paper, and for the numerous hours of conversation around personhood and possible worlds, both of which led me to the contemplations behind the strong compatibilist proposal, and to this particular manifestation of this idea.

References

- Augustine 391/2002. "Divine Foreknowledge, Evil and the Free Choice of the Will." In *Free Will*, edited by R. Kane, 259–63. Hoboken, NJ: Blackwell Publishers.
- Ayer, A. J. 1954. "Freedom and Necessity." In *Philosophical essays*, edited by A. J. Ayer, 3–20. New York, NY: St. Martin's Press.
- Chisholm, R. 1964/2004. "Human Freedom and the Self." In *Oxford Readings in Philosophy: Free Will*, edited by G. Watson, 2nd ed, 26–37. Oxford, UK: Oxford University Press.
- Crisp, T. M. 2004. "Reply to Ludlow." In *Oxford Studies in Metaphysics*, edited by D. W. Zimmerman, 37–46. Oxford, UK: Oxford University Press.
- Frankfurt, H. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66 (23): 829–39.
- Frankfurt, H. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68 (1): 5–20.
- Hobart, R. E. 1934. "Free Will as Involving Indeterminism and Inconceivable without it." *Mind* 43: 1–27.
- Hobbes, T. 1651/2018. *Leviathan*. Minneapolis, MN: Learner Publishing Group.
- Holbach, P. H. T. 1770/2005. *The system of nature or, the laws of the moral and physical world*, Vol I, (Samuel Wilkinson, Trans., Robert D. Richardson Jr. Commentator). <https://www.gutenberg.org/files/8909/8909-h/8909-h.htm> (Volume II): The Project Gutenberg. Also available at <https://www.gutenberg.org/files/8910/8910-h/8910-h.htm>.
- Hume, D. 1784/1975. *An enquiry Concerning Human Understanding*, edited by P. H. Nidditch. Oxford, UK: Clarendon Press.

- Johnson, D. K. 2009. "God, Fatalism, and Temporal Ontology." *Religious Studies* 45 (4): 435–54.
- Kane, R. 1999/2004. "Responsibility, Luck, and Chance: Reflections on Free Will and Indeterminism." In *Oxford Readings in Philosophy: Free Will*, edited by G. Watson, 2nd ed, 299–321. Oxford, UK: Oxford University Press.
- Kane, R. 2000. "The Dual Regress of Free Will and the Role of Alternative Possibilities." *Noûs* 34 (14): 57–79.
- Kripke, S. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Lenart, B. A. (2007). *Hardwired freedom: Illusion as a vehicle for moral responsibility* [Unpublished master's thesis]. University of Alberta. <https://era.library.ualberta.ca/items/3abc4c40-0b0d-4f83-9ff9-9f6b0f26f59a#.WKXmRhIrl1h>.
- Lenart, B. A. 2019. "Why we shouldn't pity Schrödinger's kitty: Revisiting David Lewis' worry about quantum immortality in a branching multiverse." *Metaphysica: International Journal for Ontology & Metaphysics* 20 (1): 117–36.
- Lewis, D. 1973. "Causation." *The Journal of Philosophy* 70 (17): 556–67.
- Lewis, D. 1976a. "Are We Free to Break the Laws?" In *Philosophical Papers*, vol. II, 291–8. Oxford University Press.
- Lewis, D. 1976b. "The Paradoxes of Time Travel." *American Philosophical Quarterly* 13 (2): 145–52.
- Lewis, D. 1979. "Counterfactual Dependence and Time's Arrow." *Noûs* 13 (4): 455–76.
- Lewis, D. 1983. "Survival and Identity." In *Philosophical Papers*, I, 55–77. Oxford, UK: New Oxford University Press.
- Lewis, D. 1986. *On the Plurality of Worlds*. Hoboken, NJ: Blackwell Publishing.
- Lewis, D. 1999. "Evil for Freedom's Sake?" In *Papers in Ethics and Social Philosophy*, 101–27. Cambridge, UK: Cambridge University Press.
- Lewis, D. 2004. "How Many Lives Has Schrödinger's Cat?" *Australasian Journal of Philosophy* 82 (1): 3–22.
- Linsky, B., and E. N. Zalta. 1994. "In Defense of the Simplest Quantified Modal Logic." *Philosophical Perspectives* 8: 431–58.
- Ludlow, P. 2004. "Presentism, Triviality, and the Varieties of Tensism." In *Oxford Studies in Metaphysics*, edited by D. W. Zimmerman, 21–36. Oxford, UK: Oxford University Press.
- McKenna, M., and D. J. Coates. 2021. "Compatibilism." In *The Stanford Encyclopedia of Philosophy*, edited by Zalta, E. N. Also available at <https://plato.stanford.edu/entries/compatibilism/>.
- Mackie, P. 2017. "Transworld Identity." In *The Stanford Encyclopedia of Philosophy*, edited by Zalta, E. N. Also available at <https://plato.stanford.edu/archives/win2017/entries/identity-transworld/>.
- McDaniel, K. 2004. "Modal Realism with Overlap." *Australasian Journal of Philosophy* 82: 137–52.
- Nozick, R. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.
- Papineau, D. 2004. "David Lewis and Schrödinger's Cat." *Australasian Journal of Philosophy* 82 (1): 153–69.
- Parfit, D. 1986. *Reasons and Persons*. Oxford, UK: Oxford University Press.
- Pereboom, D. 2001. *Living without Free Will (Cambridge Studies in Philosophy)*. Cambridge, UK: Cambridge University Press.
- Pereboom, D. 2005. "Defending Hard Incompatibilism." *Midwest Studies In Philosophy* 29: 228–47.

- Searle, J. 2004. *Mind: A Brief Introduction*. New York, NY: Oxford University Press.
- Smilansky, S. 2000. *Free will and Illusion*. New York, NY: Clarendon Press.
- Strawson, P. 1963/2004. "Freedom and Resentment." In *Oxford Readings in Philosophy: Free Will*, edited by G. Watson, 2nd ed, 72–93. Oxford, UK: Oxford University Press.
- van Inwagen, P. 1983. *An Essay on Free Will*. New York, NY: Clarendon Press.
- van Inwagen, P. 1998. "The Mystery of Metaphysical Freedom." In *Metaphysics: The Big Questions*, edited by P. van Inwagen, and D. W. Zimmerman, 365–74. Hoboken, NJ: Blackwell Publishing Ltd.
- Wegner, D. M. 2003. *The Illusion of Conscious Will*. Cambridge, MA: MIT Press.
- Yagisawa, T. 2010. *Worlds and Individuals, Possible and Otherwise*. New York, NY: Oxford University Press.
- Zagzebski, L. 2021. "Foreknowledge and Free Will." In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta Also available at <https://plato.stanford.edu/entries/free-will-foreknowledge/#CompRespTheoFata>.