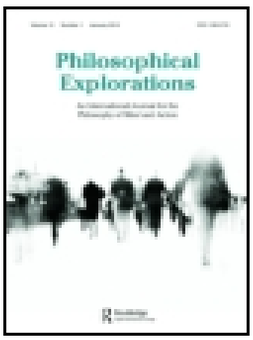

Routledge
Taylor & Francis Group

Philosophical Explorations

An International Journal for the Philosophy of Mind and Action

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/rpex20

Free will, determinism, and the right levels of description

Leonhard Menges

To cite this article: Leonhard Menges (2021): Free will, determinism, and the right levels of description, Philosophical Explorations, DOI: 10.1080/13869795.2021.1937679

To link to this article: https://doi.org/10.1080/13869795.2021.1937679

Published online: 11 Jun 2021.

PHILOSOPHICAL EXPLORATIONS
https://doi.org/10.1080/13869795.2021.1937679

Routledge
Taylor & Francis Group

Free will, determinism, and the right levels of description

Leonhard Menges

Department of Philosophy (KGW), University of Salzburg, Salzburg, Austria

ABSTRACT

Recently, many authors have argued that claims about determinism and free will are situated on different levels of description and that determinism on one level does not rule out free will on another. This paper focuses on Christian List's version of this basic idea. It will be argued for the negative thesis that List's account does not rule out the most plausible version of incompatibilism about free will and determinism and, more constructively, that a level-based approach to free will has better chances to meet skeptical challenges if it is guided by reasoning at the moral level – a level that has not been seriously considered so far by proponents of this approach.

ARTICLE HISTORY
Received 17 November 2019
Accepted 30 April 2021

KEYWORDS
Free will; Determinism; Alternative possibilities; Christian List; Levels; Responsibility

1. Introduction

A classical compatibilist account of free will says, roughly, that physical determinism and free will are situated on different levels and that determinism on one level does not threaten free will on another (see, esp., Kenny [1978] 2012, chap. 2; Dennett 2004; see also Melden 1961). For a while, this approach has been largely out of sight from the point of view of mainstream debates about free will.[1] Recently, however, it has made an impressive comeback (see, e.g. Berofsky 2012; Roskies 2012; Ismael 2013; 2016; List 2014, 2019).

The aim of this paper is to critically discuss a recent and especially interesting level-based approach developed by Christian List that he calls *compatibilist libertarianism* (2019, 9), which I will present in section 2. In particular, I will argue for two theses. First, List's account does not rule out the most plausible version of incompatibilism about free will and determinism (sections 3 and 4). Second, more generally and constructively, a level-based approach to free will has better chances to meet incompatibilist and skeptical challenges if it is guided by reasoning at the moral level – a level that has not been seriously considered so far by proponents of this approach (section 5).

Let me begin with some clarifications. First, I will assume that free will requires alternative action possibilities and I will not say anything about Frankfurt cases. Second, I assume that physical determinism is the claim that given

> the complete physical state of the world at any point in time, only one future sequence of events is physically possible. 'Physically possible,' in turn, means 'compatible with the fundamental physical laws' (List 2019, 87).

CONTACT Leonhard Menges leonhard.menges@sbg.ac.at

Finally, I stipulate that 'compatibilism' refers to the view that humans can have free will even if physical determinism is true. 'Incompatibilism' refers to the denial of compatibilism.

2. List's compatibilist libertarianism

Intuitively, it seems very plausible that we choose among different options in our daily lives – coffee over tea, to stay single rather than to marry, to become a philosopher rather than a lawyer – and that free will involves the ability to make these choices. Accepting this picture confronts us with a classical argument to the conclusion that no human has free will if the world is deterministic:

(1) Free will requires that, at least in relevant situations, agents have alternative possibilities and can do otherwise.
(2) Determinism implies that, in any situation, there are no alternative possibilities and nobody can do otherwise.
(C) Free will and determinism are incompatible.

Many compatibilists have tried to argue that (2) is false by offering interpretations of alternative possibilities or 'can do otherwise' that have room in a deterministic world. Dispositionalist accounts, for example, say, very roughly, that 'can do otherwise' should be understood as the agent's having a set of dispositions to act otherwise in relevant circumstances. Take someone's ability to speak Spanish. When she is asleep she does not speak Spanish, but she has the disposition to speak Spanish in relevant circumstances. Thus, she can do otherwise than she, in fact, does. Determinism surely does not rule out these kinds of abilities to do otherwise.

Many have doubts that this is a convincing reply to the incompatibilist challenge (see, esp., Clarke 2009). List imagines a pianist who is disposed to play a Mozart sonata flawlessly in normal circumstances, but who freezes 'under the special pressure of an audition' (List 2019, 85). According to dispositionalism, the pianist *can* play the sonata flawlessly in the relevant sense, even if she does not. But List objects that this is 'only a watered-down notion of alternative possibilities: a notion that does not capture what we normally mean by saying that someone 'could have done otherwise'' (List 2019, 85).

This very brief discussion of dispositionalism is relevant for figuring out how to evaluate List's own account. What he is looking for is an understanding of 'can do otherwise' that is compatible with physical determinism and not 'watered-down'. I propose to understand this in the following way: his interpretation of 'can do otherwise' should be close to what we have in mind in everyday thinking about alternative possibilities and clearly relevant for discussing free will in philosophical discourse. In the remainder of this section, I will present his account. Then, I will discuss whether he achieves these aims.

The basic idea of List's level-based reply to the incompatibilist argument is that it rests on the mistake of not distinguishing between two different levels of description (for more details, see List 2018, 2019, 89–90). Very generally, the view says that understanding different phenomena requires describing them on appropriately fine- or coarse-grained levels by using the corresponding concepts. Understanding the basic laws of nature, for example, requires describing them at the fundamental physical level by using

concepts like particles, forces, and so on. The evolution of species or the development of the employment rate, however, require very different concepts and higher, more coarse-grained levels of descriptions. We will not be able to understand the evolution of species by describing the behavior of individual particles. We need notions like heredity, genotype, and others that 'abstract away from the microphysical details' (List 2018, 1). Whether an agent can do otherwise must, according to this view, be described at the psychological or agency level, using concepts like belief, desire, or intention. Actions cannot be fully and informatively made sense of by using only physical concepts like forces and particles.

Based on the idea that physical determinism is a claim at the physical level and that claims about alternative action possibilities are situated on the psychological or agency level, List reformulates the premises of the incompatibilist argument (see List 2019, 87):

(1) Free will requires that, at least in relevant situations, more than one course of action be *possible for the agent*.
(2) Physical determinism implies that, given the complete physical state of the world at any point in time, only one future sequence of events is *physically possible*.

Crucially, it does not follow that physical determinism and free will are incompatible. The claim that only one course of events is physically possible is a claim at the physical level of description and silent with regard to whether or not an agent has the possibility to act otherwise, which is a claim at the psychological level. What would rule out alternative action possibilities is physical determinism in combination with what List calls the *linking thesis*: '*If*, given the complete physical state of the world at any point in time, only one future sequence of events is physically possible, *then*, in any situation, only one course of action is ever possible for an agent' (List 2019, 88 italics in original).

List then argues that the linking thesis is false. He does so by giving an account of how the physical and psychological levels relate to each other. According to him, the psychological supervenes on and can be multiply realized by the physical.[2] That is, if there is a change at the psychological level, then there must be a change at the subvenient physical level, and different patterns at the physical level, such as different arrangements of particles, can realize one and the same psychological-level state (see List 2019, 162, n. 8). For example, my desire to drink water supervenes on the physical in the sense that if the desire gets stronger this must correspond to a change on the physical level. The desire is multiply realizable in the sense that different arrangements of particles and forces can realize the very same desire.

List argues that this account of how the physical and the psychological relate to each other opens the door for different action possibilities even if physical determinism is true. Take Figure 1, a simplified version of List's own illustration (2019, 94). The whole figure represents an agent and the agent's environment, say, me in my office. Rows stand for the time periods t1 to t3, rectangles for possible psychological states, dots for possible physical states and lines for possible histories, that is, roughly, how my environment and I could develop from time t1 to t3. Take the rectangle at t1 to stand for my desire to drink water. This desire is multiply realizable in the sense that it can be realized by two different physical states (two dots in one rectangle). Now focus on the physical level, that is, the dots. There is no branching in the histories. If my environment and I

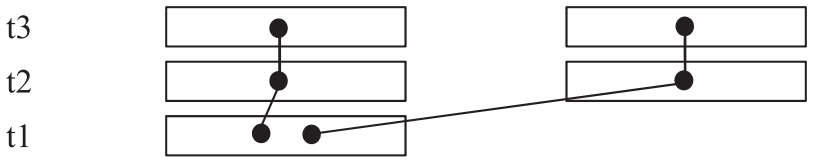

Figure 1.

are in a certain physical state, only one future sequence of states is physically possible. That is, at the physical level, the history from t1 to t3 is deterministic.

Now consider Figure 2 which, like Figure 1, represents me and my environment but, unlike Figure 1, only focuses on the psychological level (the rectangles). There is branching in the histories. My desire to drink water at t1 can be followed by the psychological states on the left side, say, my intention at t3 to get up and get a glass of water, and it can be followed by the psychological states on the right side, say, my intention at t3 to not get up, finish the paragraph, and get a glass of water afterwards. Thus, my being in one psychological state at t1 is consistent with being in two different psychological states at t2 and t3. That is, at the psychological level, my history is indeterministic.

This illustrates how, even if physical determinism is true, different future psychological states are possible for the agent. As it is very plausible that these different psychological states, including different intentions, desires, and so on, lead to different actions, it follows that different future actions are possible for the agent. That is, at t1, it is possible for me to get up at t3 and not to get up at t3. List concludes that 'determinism at the physical level can coexist with indeterminism at the level of agency' (List 2019, 95). Therefore, the linking thesis is false, the incompatibilist argument is refuted, and alternative action possibilities are compatible with physical determinism.

In the following two sections, I will discuss this line of reasoning and argue that List does not fully meet his goals to offer a compatibilist interpretation of 'can do otherwise' that is close to everyday thinking (Section 3) and clearly relevant for free will (Section 4).[3] In order to do so, I will put aside many issues that some may find worthy of discussion. For example, I will not scrutinize List's notions of multiple realizability and supervenience and whether they deliver a non-reductive account of the mental. Similarly, I will put aside the question of how multiply realizable mental states can be causally efficacious, as List claims.[4] I will also accept some key elements of List's level-based approach, namely that, first, different phenomena need to be described at different levels of description; second, determinism, as it is typically understood, is a claim on the physical level; and third, the claim that agents can do otherwise is a claim on the psychological level. Granting this, I will argue that List's account, nonetheless, does not achieve its main goals.

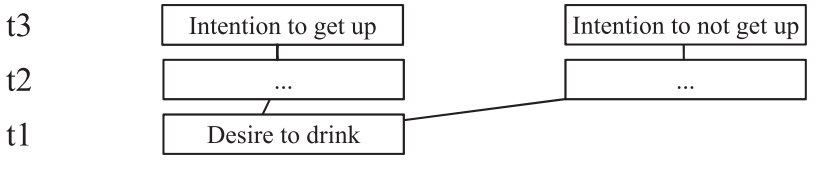

Figure 2.

3. Can and can't do otherwise

In this section I will argue that there is reason to doubt that List's compatibilist interpretation of 'can do otherwise' is the one at issue in everyday thinking. My starting point is the question of whether the truth of physical determinism has implications for the psychological level. The linking thesis says: Yes, physical determinism leads to psychological determinism. List replies: As the psychological supervenes on and is multiply realizable by the physical, there can be physical determinism and psychological indeterminism. In what follows, I will first present an everyday sense of 'can do otherwise' that would be ruled out by physical determinism. Then I will defend this line of reasoning against possible objections.

Imagine Afia who, at t1, wants to drink water. Imagine also that Figures 1 and 2 correctly depict her and her environment and possible histories from t1 to t3. Thus, we know the psychological states she can be in (rectangles) and we know the physical states that can realize these psychological states (dots). So far, this is not new. But let us imagine that we know one more thing. After t3 we learn that Afia in fact intended to get up at t3 and did. As we know that there was only one trajectory from her physical and psychological states at t1 to her intention to get up at t3, we can infer which physical state realized her desire to drink at t1.[5] In short, after t3 we know that Figures 3 and 4 were true of Afia, where the thicker lines and larger larger dots represent the actual history and the thinner lines and smaller dots reprsent the possible but not-actual history.

According to List's account, different courses of actions were possible for Afia at t1. Her desire to drink *could have been* realized by the physical states represented by the large and small dots which were, therefore, consistent with her desire. If Afia's desire at t1 *would have been* realized by the state represented by the small dot, she would have developed other psychological states and would have acted differently. However, *in fact*, her psychological state was realized by the physical state represented by the large dot, which yielded the sequence which realized her getting up at t3.

Now let us ask: Could Afia at t1 choose between getting up and not getting up at t3? It seems natural to reply along the following lines: 'Contrary to what Afia believed, there was nothing she could have done in order to not get up at t3' or 'She did not believe this, but, in fact, there was, at t1, only one course of action possible for Afia'. Indeed, these claims seem to be true. Other courses of action were *consistent* with her desire to drink at t1, but there is a sense in which she *could not* have performed them. Now someone may ask why we believe this: what explains that, at t1, only one course of action was possible for Afia? The most natural answer is that the fact that her history from t1 to t3 was determined at the physical level explains that she could not have done otherwise.

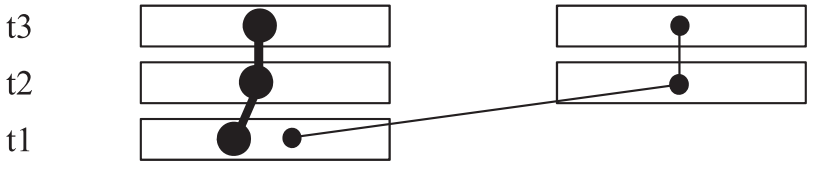

Figure 3.

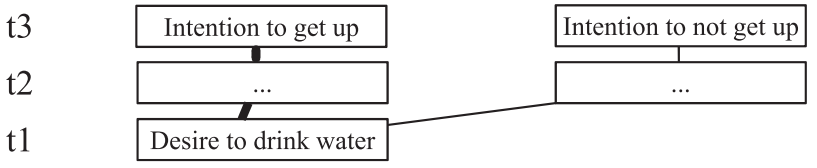

Figure 4.

The case of Afia highlights a common idea that was covered before, namely that each psychological state is *in fact* realized by a specific physical state even if it *can* be realized by several different ones. Assuming that the realizing physical state is part of a deterministic history, it is very plausible that the realized psychological states follow this deterministic history. Thus, the case of Afia is meant to support the claim that physical determinism leads to a form of psychological determinism (Gebharter 2020, sec. 3 makes a similar point).

As an intermediate result, Afia could have done otherwise in the sense identified by List even if physical determinism is true. However, the case supports the idea that there is an important everyday understanding of 'can do otherwise' which is such that Afia could not have done otherwise in this sense, which is explained by the fact that her history at the physical level was deterministic. Thus, the line of thinking presented so far supports a moderate version of the linking thesis: if, given the complete physical state of the world at any point in time, only one future sequence of events is physically possible, then, in one everyday sense of 'can do otherwise', in any situation, agents cannot do otherwise than they in fact do.

Let me discuss some objections against this line of reasoning.[6] First, one could object that the discussion of Afia's case involves an unrealistic picture of what we know. Science is extremely far away from what we imagine to know in the case of Afia. However, it is important for List that his account of free will is scientifically acceptable.

As a first reply, the scenario of Afia does not seem to be much further away from what science tells us than what List imagines in his line of reasoning. When arguing against the linking thesis, he depicts a more complex system than we imagined above (see List 2019, 92–97). When looking at this system, we also know the possible psychological states of the agent and the physical states that can be their realizers. The only difference between List's case and the one involving Afia is that, in the latter, we imagine that we learn after t3 which psychological and physical state Afia in fact was in at t3. This allows us to infer her history at the physical and psychological level, which is also possible in List's model, even if he does not explicitly say it. Thus, it is true that the reasoning presented here involves imagining that we have more knowledge than List's line of reasoning requires us to imagine. However, I see no reason to think that the line between a scientifically acceptable imagined scenario and a scientifically unacceptable imagined scenario runs exactly between us.

As a second reply, assume that the line of reasoning that involves Afia's case is much further away from what science tells us than List's. Would this be a problem? I do not see that it would. Philosophy is full of thought experiments in which we imagine being in unrealistic situations and in which we ask ourselves something like 'would life in the experience machine be best?' (see Moore 2019) or 'does Mary learn some new

information?' (see Nida-Rümelin and O Conaill 2019). These lines of thinking are meant to tell us something interesting about the good life and about what kind of information there is. Arguing in this way does not *use* the empirical sciences, but I see no reason to think that it is scientifically *unacceptable*. And the same holds when we ask ourselves if Afia could have done otherwise in order to learn something about 'can do otherwise'. Therefore, I conclude that this objection on behalf of List is not convincing.

A second objection says that the discussion of Afia begs the question against List's view because it focuses too much on the physical level. List argues that the ability to do otherwise is a psychological phenomenon that needs to be described at the psychological level. But the way I described Afia's case above, the objection goes, focuses on the physical states and infers physical-level trajectories. It is no surprise that we cannot find an ability to act otherwise on this physical level, the objection says. But this does not speak against List's view that there is an ability to act otherwise at the psychological level.

As a response, recall the dialectical situation. The question at hand is whether determinism at the physical level has implications for the psychological level. Thus, at this point we cannot assume that the psychological is very independent from the physical. Whether this is so is the question at issue. List gives an answer to this question by presenting a line of thinking illustrated by Figures 1 and 2. His conclusion is, roughly: the psychological is largely independent from the physical. The line of reasoning that involves Afia's case supports a different answer to the question, namely: the psychological is more dependent on the physical than List claims. At this point, there is an exchange of arguments, but no question-begging.

Now recall the structure of the argument that involves Afia's case. It works by, first, trying to trigger the intuition that, in one important sense, Afia cannot do otherwise, which is a claim on the psychological level. The argument says, second, that the best explanation for this is that Afia's history is deterministic at the physical level. Again, I do not see that this begs the question against List. Rather, this is an independent argument to the conclusion that physical determinism rules out one everyday sense of 'can do otherwise'.

The third objection will provide a bridge to the next section. The objection concedes that Afia lacks the ability to do otherwise in some sense. But, the objection says, this sense of 'can do otherwise' is not the one at issue in everyday thinking about agency. Let me elaborate.

List modifies a case from Angelika Kratzer in which a Vienna Circle philosopher answers a judge's question of whether a murderer could have done otherwise. The philosopher claims that, 'given the whole microphysical history leading up to the crime', including 'all relevant neurophysiological processes leading up to the murderer's action', the murderer could not have done otherwise than she did (List 2019, 106). The point of the case is that the philosopher misunderstood the judge's question. While the judge wanted to know something like 'Could the murderer have acted otherwise given our best psychological theories?', the philosopher understood the judge as asking 'Could she have acted otherwise given our best physical theories?' The objection says that our judgment that Afia cannot do otherwise is like the Vienna Circle philosopher's judgment about the murderer: it relies on a misunderstanding of how we use 'can' in everyday life.

As a first reply, it seems plausible that the Vienna Circle philosopher misunderstood the judge because the judge is *a judge*. This suggests that she is interested in an

understanding of 'can do otherwise' that is legally relevant. Facts about the microphysical history of an action, however, are not relevant in the legal systems I know. But I see no reason to hold that everyday thinking about 'can do otherwise' is restricted to an understanding of 'can do otherwise' that is legally relevant. Therefore, it would be better to have a non-legal example.

Imagine that Afia wins the Nobel Prize for physics. There was a moment in her career in which she thought about becoming a lawyer or a physicist. The history that led to her becoming a physicist looks like Figures 3 and 4. That is, the thick line went to her intention to become a physicist and the thin line through the intention to become a lawyer. After she won the Nobel Prize, someone asks us if Afia could have become a lawyer.

This question can be understood in different ways. Here are three options. First, the person wants to know if Afia has the intellectual and psychological dispositions and virtues that a lawyer needs – this is an understanding of 'can do otherwise' close to the dispositionalist understanding sketched in section 2. In that case, we do not know the answer. Second, the person wants to know whether Afia's becoming a lawyer was consistent with some of her earlier psychological states – this is, roughly, List's understanding of 'can do otherwise'. The answer would be 'Yes'. Third, the person wants to know whether Afia had access to a course of physical and psychological events that would have had the result that she ended up being a lawyer – this is the interpretation presented in this section. The answer would be 'No'.

Is one of these interpretations more natural than the others? Personally, I find that the first (dispositionalist) interpretation is the one that comes to mind most immediately in this case. But with regard to the second and third interpretation of 'can do otherwise', I see no difference in terms of what is more natural in everyday thinking. In particular, I see no reason to hold that the sense of 'can do otherwise' which focuses on the physical and psychological accessibility of different actions is further away from everyday thinking than a sense of 'can do otherwise' that focuses on consistency with earlier psychological states. Thus, I see no reason to believe that the judgment that Afia cannot do otherwise rests on a misunderstanding of the everyday notion of 'can'.

But there is still an open question. The discussion so far was meant to support the idea that if physical determinism is true, agents lack the ability to do otherwise in one important sense. But, as the case highlighted again, agents may still have the ability to do otherwise in other important senses. Which of these senses is most relevant for free will? This will be the guiding question of the next section.[7]

To sum up, one of List's main goals is to offer an account of 'can do otherwise' that is compatibilist and close to everyday thinking. I have argued for two claims in this section. First, there is an important everyday sense of 'can do otherwise' that is incompatibilist; and second, it is far from clear that the sense List picks out is the one which is most relevant in our everyday thinking about agency.

4. Free will and responsibility

List's second main goal identified above is to offer a compatibilist account of 'can do otherwise' which is clearly relevant for free will. In order to find out if he achieves it, I will first present List's most important philosophical opponents, namely free will skeptics

– or skeptics, for short. The skeptics' position will, then, serve as a foil in order to test if List's account helps in meeting the skeptical challenge. I will argue that it does not.

A core motivation for thinking about free will is a concern about responsibility and, especially, our practices of holding people responsible by blaming, praising, and being grateful towards them. Many authors, including List, assume 'the presence of free will somewhere along the relevant chain of events to be a *necessary condition* for a salient form of moral responsibility' (List 2019, 28 italics in original; see also Mele 2006; Levy 2011; Vargas 2013; Pereboom 2014). Thus, agents can perform actions out of free will without being responsible for them, perhaps because they lack a certain knowledge. But if they are responsible for their actions, then they performed them out of free will.

A standard, broadly Strawsonian account of responsibility says that being responsible should be understood in terms of being an appropriate target of certain responses. In what follows, I will focus on responsibility for objectionable actions. The idea is that an agent is responsible for an objectionable action just in case it is appropriate to blame the person for it (see, e.g. Wallace 1994; Watson [1996] 2004; Darwall 2006; Schulte 2014; Pereboom 2014; Shoemaker 2015; Menges forthcoming).

The appropriateness at issue can be understood in different ways, but most skeptics spell it out in moral terms, namely in terms of justice, fairness, or desert (see Strawson 1994; Levy 2011; Waller 2011; Caruso 2012; Pereboom 2014; for an overview, see Caruso 2018). Note that there is an important difference between the claim that agents are responsible for objectionable actions in the sense that it is fair, just, or deserved to blame them for the actions and the claim that it is all-things-considered justified, permissible, or good to do so. It can be fair, deserved, or just to blame agents for some action without its being fully justified to blame them (e.g. if blaming them has disastrous consequences) and the other way around (e.g. if blaming them has fantastic consequences). Moreover, claiming that agents are responsible for bad actions in the sense that blaming them would be fair, just, or deserved does not necessarily commit one to the retributive idea that harming them would be noninstrumentally good. Alternatively, claims about fair, deserved, or just blame can be understood as implying that there is a conditional reason to blame the agents (see Nelkin 2016, sec. 2), that the agents' interests in not being blamed can be discounted in consequentialist calculations (see Levy 2011, 3), or that the blamed agent owes the victim an apology (see Pettit 2007, sec. 1).[8]

On this rough picture, free will is an important necessary condition for a person's being responsible for an objectionable action in the sense that blaming the person for it would be just, fair, or deserved, while this kind of appropriateness does not imply that blaming is all-things-considered good or justified.[9] Free will skeptics contend that there is very good reason to doubt that humans have free will in this sense. Then, the challenge from determinism is whether it can be, in the relevant sense, appropriate to blame humans in a deterministic world. If it cannot, we have strong moral reason to reform our blame practices if we find out that physical determinism is true.

Some authors reply that they are not concerned with this kind of free will (see, e.g. Berofsky 2012, chap. 1). List, however, seems to accept such a picture of free will and responsibility (see List and Pettit 2011, sec. 7.1; List 2019, 154; List's contribution to Caruso, List, and Clark 2020).[10] Moreover, even if some defenders of free will are not interested in the relation between free will and the moral appropriateness of the responsibility practices, they should be interested in it if they want to engage with their strongest philosophical

opponents, namely the free will skeptics I sketched above. Otherwise, they would simply talk past the skeptics.

The picture of responsibility I have just presented suggests a way to test whether the kind of alternative possibilities List picks out and shows to be compatible with physical determinism is relevant for free will discourse. His account of 'can do otherwise' should help to explain when it is and when it is not appropriate to blame agents. The best way to test this is to imagine two agents who act morally objectionably, fulfill all epistemic and other agency conditions that are necessary for being responsible, and only differ in one relevant respect: one of them can do otherwise in List's sense and the other cannot. If List's account is relevant for free will skepticism, then the first is an appropriate target of blame, while the second is not.

Imagine that Charlie is a great physicist, but not as great as Afia. He fulfills all the relevant epistemic conditions for moral responsibility, his agency is fully intact, and so on. At t1 Charlie is in the psychological state of desiring to win the Nobel Prize before Afia does. This desire is realizable by two different physical states and consistent with two different histories at the psychological level from t1 to t3 (see Figures 5 and 6). One of the psychological histories involves the beliefs that Afia will probably win the next Nobel Prize and that the only way to prevent this is to kill her at t3, his doubts about whether this is a good idea, and then finally the intention to kill Afia. The other history ends with the intention to not kill Afia. Thus, according to List's interpretation, Charlie can, at t1, choose between killing Afia at t3 and not killing Afia at t3. As we assume that he fulfills all other conditions for responsibility, List's account implies that Charlie would be responsible for what he intends to do at t3. Then, Charlie kills Afia at t3, such that he is, on List's account, an appropriate target of blame for it.

Now consider Dadi in a world very close to the world in which Charlie kills Afia and who differs from Charlie in only one relevant way. Namely, without any fault of his own, Dadi's psychological states between t1 and t3 are *not* multiply realizable. That is, each of his psychological states can only be realized by one physical state (see Figure 7). We could imagine that, in Dadi's world, mental states in general are not multiply realizable (reductionists claim that our world is like that; for an overview see Bickle 2016). Or we can imagine that a very special constellation of events for which Dadi is not responsible made it the case that, between t1 and t3, Dadi's psychological states are not multiply realizable. Then, the sequence of physical and psychological events is identical with the sequence described above with regard to Charlie. Dadi believes that Afia will probably win the next Nobel Prize, that the only way to prevent this is to kill her at t3, he doubts whether this is a good idea, and then he forms the intention to kill Afia and actually does it (see Figure 8).

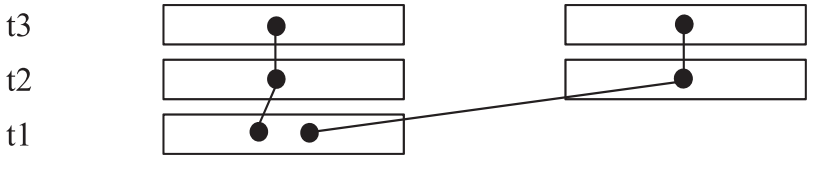

Figure 5.

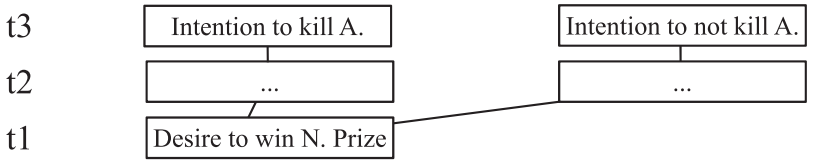

Figure 6.

In contrast to Charlie, Dadi's psychological states are not multiply realizable and, therefore, not consistent with physical and psychological states other than the ones he is in fact in. Thus, in List's sense, Charlie can, but Dadi cannot do otherwise.

If List's account of 'can do otherwise' is relevant for the free will debate as I described it above, then the absence of other physical and psychological states that are consistent with Dadi's desire to win the Nobel Prize before Afia explains why he is, in contrast to Charlie, not responsible for killing Afia. Then, it would be morally appropriate to blame Charlie for killing Afia, but it would be inappropriate to blame Dadi even though he committed the same deed.

I find this implausible. If one assumes that it is morally appropriate to blame Charlie, why should one think that it is inappropriate to blame Dadi?

Note, first, how similar they are. Their actual psychological histories were almost identical: they had the same quality of will, first- and second-order desires, the very same line of thinking led to their killing Afia, and so on.[11] Thus, those who think that some of these factors (together with the other relevant agency and epistemic properties we assume them to have) are sufficient for responsibility will take Charlie and Dadi to be equally (in)appropriate targets of blame.

The strongest objection against the claim that Charlie and Dadi differ with respect to morally appropriate blame is the following: a general and very plausible principle says that if it is morally appropriate (fair, just, or deserved) to blame one agent but not another, then there must be a morally relevant difference between them. Applied to our case, this principle demands that, if it is morally appropriate to blame Charlie and inappropriate to blame Dadi, then there must be a morally relevant difference between them. List's view suggests two options to account for this difference. The first is to say that the fact that Charlie's mental states are multiply realizable while Dadi's mental states are not is morally relevant. The second option says that the fact that Charlie's psychological state at t1 is consistent with different future courses of actions, while Dadi's psychological state at t1 is not, is morally relevant. However, both claims are quite implausible. They are neither intuitively convincing nor do I see arguments for why one of them should be true.[12] It seems much more plausible to say that there is

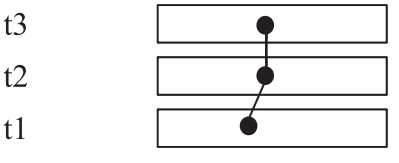

Figure 7.

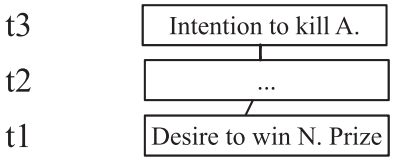

Figure 8.

no morally relevant difference between Charlie and Dadi, such that there is no difference between them with regard to fair, just, or deserved blame.[13]

Recall how we got here. In the preceding section, I argued that, if determinism is true, then an agent who can do otherwise in List's sense cannot do otherwise in another important sense. The question in this section is whether List picks out a sense of 'can do otherwise' that is relevant for philosophical debates about free will. If List's notion of 'can do otherwise' is supposed to be the same that is at issue in the debate about philosophical free will skepticism, then his account should help to explain when it is and when it is not appropriate to blame agents. However, the cases above suggest that it does not help with this issue. Two agents who only differ in the respect that one of them can while the other can't do otherwise in List's sense do not differ with regard to morally appropriate blame. If this is correct, then List picks out a sense of alternative possibilities that does not capture what is important to List's strongest philosophical opponents, namely free will skeptics. These skeptics would say about List's account what he says about dispositionalism, namely that it delivers 'only a watered-down notion of alternative possibilities: a notion that does not capture what we normally mean by saying that someone 'could have done otherwise'' (List 2019, 85).

5. Starting from the moral level of description

List's diagnosis of the skeptical challenges to free will in general and the deterministic challenge in particular is that they rely on the illegitimate blending of the physical and the psychological level (see List 2019, 150). Those who look for free will at the physical level will not find it because it is a psychological, not a physical phenomenon. Once we look at the psychological level, List says, we will find it. I have argued above that this is too quick. Adopting this approach may arguably identify *some* kind of freedom which is compatible with determinism. But it looks as if this kind of freedom is not the one skeptics are skeptical about and we have reason to worry about. In this final section, I will propose how proponents of the level-based approach can make sure that they identify the relevant kind of freedom. The idea is, roughly, that one should start thinking about free will at the moral level and then look for it at the psychological level (the approach is inspired by Wallace 1994; see also Schulte 2014). I will first sketch a rough understanding of the moral level of description. Then, I will present how thinking on the moral level helps to clarify what free will is, if it exists at all. Finally, I will compare this approach to List's.

While List mostly focuses on different levels of the empirical realm, it seems fruitful to apply the idea of levels to the non-empirical realm, too. One could argue that there is a fundamental normative level that, depending on one's metaethical commitments,

operates with notions such as *values*, *reasons*, or *obligations* (List 2018, sec. 4.4 describes the normative in terms of 'may' and 'ought'). Higher normative levels would be, for example, epistemology, ethics/morality, or political theory. These different levels operate with their own normative notions such as *knowledge* and *evidence* in epistemology, *permissibility* and *desert* in ethics or *just distribution* and *freedom* in political theory. While it seems plausible that, for example, epistemological and moral facts supervene on deeper normative facts about reasons, values, or obligations, moral facts do not supervene on epistemological facts and *vice versa* (see List 2018, sec. 4.1 for a similar discussion of geology and biology). As the question of how the normative levels relate to the physical level touches on some of the most complicated metaethical issues, I will have to leave it aside (but see List 2018, sec. 4.4).

My main proposal is that adopting the moral perspective will tell us what free will is, if it exists at all, by identifying the roles that it is expected to play. Therefore, thinking about free will should start at this level. Here is a sketch of how to do this.

Consider the skeptics again who argue that humans do not have the kind of control that is necessary for its being morally appropriate to blame them. The question of what this kind of control is needs to be discussed at the moral level by using notions like fairness, desert, wronging, claims, and so on. The question has the structure 'What is the *X*-condition which is necessary for its being morally appropriate to *y*?' Other questions with this structure are, for example, 'What kind of threat is necessary for its being legitimate to kill someone in self-defense?' or 'What must experimental subjects understand about a study for its being permissible to use their data?' These are first-order moral questions.

Identifying the control which is necessary for its being morally appropriate to blame a person involves getting clear about the nature of the relevant kind of blame. Do we want to know when it is morally appropriate to punish people, to aggressively confront them, to have anger emotions towards them, or something different (for some recent proposals, see Carlsson 2017; Menges 2017; 2020; Caruso and Morris 2017; McKenna 2019)? This question is important because different responses may require different kinds of control. For example, it seems plausible that we need more control for its being morally appropriate to physically punish us than for its being appropriate to feel resentment towards us.

Once we have made clear what kind of response we are interested in, we need a better understanding of the relevant kind of appropriateness. As I have indicated in the preceding section, 'appropriateness' can be spelled out in terms of, for example, noninstrumental goodness of harm, conditional reasons to blame, or forfeiting claims against blame (see, e.g. Levy 2011; Nelkin 2016; McKenna 2019). Identifying the relevant kind of appropriateness is relevant for free will because different kinds of appropriateness may require different kinds of control. One may argue, for example, that the kind of appropriateness which involves the non-instrumental goodness of harm requires a stronger kind of control than many other kinds of moral appropriateness.

When we have identified the kind of response and the kind of appropriateness we are interested in, we should ask what kind of control is necessary for its being in this sense appropriate to respond in this way to an agent's conduct. Must the agent be the agent-cause of the relevant action, is it enough that the agent is reason-responsive to a certain degree when she acts, that alternative actions are consistent with the agent's

earlier psychological states, must the action express the agent's deep self, or is some other kind of control necessary? Once we have answered this question, we will have identified the kind of free will the skeptic is skeptical about. All this takes place on the moral level.

Now comes the second step in the attempt to deal with the skeptical challenges: one needs to show that at least some humans have this kind of free will. This is a psychological endeavor. Armchair philosophy may succeed in making it more or less plausible that humans have a certain agential property. Folk psychological reasoning will also suggest whether or not this is so. But if we want to be *certain*, we will need our best scientific methods to prove it. To illustrate, assume the following rough picture (see, e.g. Arpaly and Schroeder 2014, chap. 7; Shoemaker 2015, pt. 1; Sripada 2016): first, the relevant kind of control over an action consists in the action's being expressive of the agent's deep self; second, the deep self consists in a sub-class of the agent's conative states and/or value judgments; third, the expression relation is a specific causal relation. Now, whether some humans have the relevant conative states and/or make the relevant value judgments and whether a certain action expresses these states and/or judgments are empirical questions that must be answered at the psychological level.

Let me compare this approach with List's. It agrees with List that the psychological, or, more generally, 'the special science perspective is the only perspective from which we are likely to be able to defend free will' (List 2014, 173). Moreover, the picture agrees that claims at the physical level about fundamental laws of nature are only relevant insofar as they are connected to the psychological level. However, from the perspective of the approach sketched here, List's account puts the cart before the horse. It does not start with a description of the roles that free will is expected to play at the moral level, but with a description of alternative action possibilities at the psychological level. Unsurprisingly, this approach fails to identify the kind of freedom skeptics worry about.

To be clear, the approach sketched here does *not* claim that reasoning at the moral level will tell us if we have free will. However, to put the proposal in a nutshell, moral reasoning should identify what we are looking for and psychological reasoning should look for it.

To sum up, differentiating between a physical and a psychological level of description helps to make clear if and how determinism and other claims about the fundamental structure of the world can be relevant for free will. However, I have argued that merely focusing on the psychological level runs the risk of not identifying the kind of free will we have most reason to worry about. Therefore, such an account should be guided by reasoning at the moral level. A detailed description of the roles that free will is expected to play at the moral level will make sure that what we are looking for at the psychological level will be relevant for the deepest worries about free will and an important part of our everyday thinking. Whether we have this kind of free will is still an open question.

Notes

1. Note, e.g. that this view is not discussed in the SEP articles on free will (O'Connor and Franklin 2018) and compatibilism (McKenna and Coates 2015).
2. List (2019, chap. 3; see also 2018, sec. 4.2) argues in some detail that phenomena at the psychological level such as beliefs, desires, or intentions cannot be fully explained in terms of or reduced to phenomena at lower levels such as the neurological or physical level.

3. See Elzein and Pernu (2017) and Gebharter (2020) for other critical discussions of List's account.
4. For details, see the account in (List and Menzies 2017; List 2019, chap. 5). The basic idea is that we should understand causation in terms of difference making, and that, thus understood, an agent's mental states often are the causes of the agent's actions, while the underlying physical states are not.
5. Note that we can imagine that we know all this without assuming that Afia's desires, beliefs, and other psychological states are reducible to the physical states she is in. The case of Afia is compatible with List's (2019, 64–69) thesis that these psychological states have the feature of intentionality that physical states do not.
6. I am very grateful to an anonymous referee for the following objections and thoughtful comments on my replies.
7. There is another question that, due to limited space, I cannot discuss in more detail. List offers an intriguing argument for the claim that humans have the ability to do otherwise and one may wonder how Afia's case relates to this argument. The very rough idea of the argument is that our best psychological theories imply that humans can do otherwise and, generally, if our best theories of a certain domain imply that something exists, then there is very good reason to think that it exists (List 2019, 74–77, 97–103). This line of reasoning does not come in conflict with what I have argued for in this section. My thesis is that if physical determinism is true, then humans lack the ability to do otherwise *in one important sense* of 'can do otherwise'. But I see no reason to think that our best psychological theories imply that humans can do otherwise in this specific sense. Perhaps these theories imply that different courses of action are consistent with some of the agents' earlier psychological states (as List says). Then, the argument would be consistent with my main thesis. See also note 10 below.
8. List (in Caruso, List, and Clark 2020) rejects a retributive justification of punishment. But it should have become clear that the claim that agents are responsible in the sense that they deserve blame does not imply that punishing them would be justified.
9. Note that this picture does not commit one to saying that the notion of free will is a moral or normative one (which is an idea that List rejects; see his contribution to Caruso, List, and Clark 2020). The view says that we can characterize free will by saying that it is a kind of control that an agent must have over an action for its being appropriate to blame the agent for it. Now assume that being an agent-cause is the strongest kind of control that is necessary for being an appropriate target of blame. Then we can analyze having free will in terms of being an agent-cause, which is, plausibly, a non-normative analysis.
10. Even though List accepts the picture of responsibility and free will I have sketched here, he does not seem to take the relation between free will and responsibility to be crucial for his theorizing about 'can do otherwise'. As I said in note 7, he aims at showing that our best psychological theories imply that humans can do otherwise in a certain sense. And he takes this sense to be relevant for free will. But skeptics do not need to accept this last assumption. They are skeptical about the kind of agency that is necessary for fair, just, or deserved blame. In order to show that the kind of freedom that List defends is relevant for the skeptical worries, List would need to show that the sense of 'can do otherwise' that is supported by psychological theories is relevant for fair, just, or deserved blame. So far, I see no reason why it should be.
11. Strictly speaking, Charlie's and Dadi's mental states are not the same because Charlie's mental states are multiply realizable, while Dadi's aren't. Thus, the precise but unnecessarily complicated formulation would say that they have exactly the same quality of will, desires, and so on, apart from the fact that they differ with regard to the multiple realizability of their mental states.
12. One may reply on behalf of List, that the fact that Charlie's psychological state at t1 is consistent with different future courses of action, while Dadi's psychological state at t1 is not, is important. It shows that Charlie, the reply goes, has a certain kind of choice, while Dadi does not. And having this kind of choice can (partly) explain why it is appropriate to

blame the chooser (thanks to an anonymous referee for pressing me on this). By contrast, the line of reasoning presented in the main text is meant to trigger the intuition that having this kind of choice is not morally relevant (in this sense). And it seems to me that the burden of proof lies on the side of List here. His position is committed to the claim that having a choice in the sense that different courses of action are consistent with one's psychological state has the normative weight that can explain deserved, fair, or just blame. But why should this be so? Why should the mere consistency of future actions be morally relevant? As I said, I do not see an intuitive answer nor do I see arguments for why this consistency is morally relevant. Therefore, I suggest that the burden of proof lies on the side of List.

13. The case of Dadi highlights an interesting implication of List's view. While most skeptics about free will think that determinism and luck undermine free will, List's account suggests that one of the most serious threats to free will is the possible falsehood of the claim that mental states are multiply realizable. That is, if it turns out that psychological states are not multiply realizable, then every human is like Dadi, no human has alternative possibilities, and, therefore, free will in List's sense would be an illusion. I find this noteworthy. In one place, List suggests that we should respond to the 'discovery' that physical determinism is true by seeing it as '"just" an interesting development in science' (List 2019, 158). Personally, I would find it much easier to see the discovery that mental states are not multiply realizable as just an interesting development in science.

Acknowledgment

I'm grateful to Hannah Altenhenger, Jan Willem Wieland, two anonymous referees, and participants of the workshop 'Responsibility and Agency', September 2019 in Salzburg, for helpful comments on earlier versions of the paper. Special thanks to one of the referees for exceptionally helpful and fair comments. Thanks to Claire Davis for proofreading.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Note on Contributor

Leonhard Menges is assistant professor for Practical Philosophy at the Department of Philosophy (KGW) at the University of Salzburg, Austria. He currently works on responsibility, blame, and the right to privacy.

References

Arpaly, Nomy, and Timothy Schroeder. 2014. *In Praise of Desire*. New York: Oxford University Press.

Berofsky, Bernard. 2012. *Nature's Challenge to Free Will*. New York: Oxford University Press.

Bickle, John. 2016. "Multiple Realizability." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2016. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2016/entries/multiple-realizability/.

Carlsson, Andreas Brekke. 2017. "Blameworthiness as Deserved Guilt." *The Journal of Ethics* 21 (1): 89–115.

Caruso, Gregg D. 2012. *Free Will and Consciousness: A Determinist Account of the Illusion of Free Will*. Lanham, MD: Lexington Books.

Caruso, Gregg D. 2018. "Skepticism About Moral Responsibility." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2018. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2018/entries/skepticism-moral-responsibility/.

Caruso, Gregg D., Christian List, and Cory J. Clark. 2020. "Free Will: Real or Illusion – A Debate." *The Philosopher* 108 (1): 61–80.
Caruso, Gregg D., and Stephen G. Morris. 2017. "Compatibilism and Retributivist Desert Moral Responsibility: On What Is of Central Philosophical and Practical Importance." *Erkenntnis* 82 (4): 837–855. https://doi.org/10.1007/s10670-016-9846-2.
Clarke, Randolph. 2009. "Dispositions, Abilities to Act, and Free Will: The New Dispositionalism." *Mind; A Quarterly Review of Psychology and Philosophy* 118 (470): 323–351.
Darwall, Stephen. 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge, MA: Harvard University Press.
Dennett, Daniel C. 2004. *Freedom Evolves*. New York: Penguin.
Elzein, Nadine, and Tuomas K. Pernu. 2017. "Supervenient Freedom and the Free Will Deadlock." *Disputatio* 9 (45): 219–243. https://doi.org/10.1515/disp-2017-0005.
Gebharter, Alexander. 2020. "Free Will as a Higher-Level Phenomenon?" *Thought: A Journal of Philosophy* 9 (3): 177–187. https://doi.org/10.1002/tht3.461.
Ismael, J. T. 2013. "Causation, Free Will, and Naturalism." In *Scientific Metaphysics*, edited by Don Ross, James Ladyman, and Harold Kincaid, 208–235. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199696499.003.0010.
Ismael, J. T. 2016. *How Physics Makes Us Free*. New York: Oxford University Press.
Kenny, Anthony. (1978) 2012. *Freewill and Responsibility*. London: Routledge.
Levy, Neil. 2011. *Hard Luck: How Luck Undermines Free Will and Moral Responsibility*. Oxford: Oxford University Press.
List, Christian. 2014. "Free Will, Determinism, and the Possibility of Doing Otherwise." *Noûs* 48 (1): 156–178. https://doi.org/10.1111/nous.12019.
List, Christian. 2018. "Levels: Descriptive, Explanatory, and Ontological." *Noûs*. Advance online publication. https://doi.org/10.1111/nous.12241.
List, Christian. 2019. *Why Free Will Is Real*. Cambridge, MA: Harvard University Press.
List, Christian, and Peter Menzies. 2017. "My Brain Made Me Do It: The Exclusion Argument Against Free Will, and What's Wrong with It." In *Making a Difference: Essays on the Philosophy of Causation*, edited by Helen Beebee, Christopher Hitchcock, and Huw Price, 269–285. New York: Oxford University Press.
List, Christian, and Philip Pettit. 2011. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. New York: Oxford University Press.
McKenna, Michael. 2019. "Basically Deserved Blame and Its Value." *Journal of Ethics and Social Philosophy* 15 (3). https://doi.org/10.26556/jesp.v15i3.547.
McKenna, Michael, and D. Justin Coates. 2015. "Compatibilism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2015. http://plato.stanford.edu/archives/sum2015/entries/compatibilism/.
Melden, Abraham I. 1961. *Free Action*. London: Routledge & Kegan Paul Limited.
Mele, Alfred R. 2006. *Free Will and Luck*. New York: Oxford University Press.
Menges, Leonhard. 2017. "The Emotion Account of Blame." *Philosophical Studies* 174 (1): 257–273.
Menges, Leonhard. 2020. "Blame It on Disappointment: A Problem for Skepticism About Angry Blame." *Public Affairs Quarterly* 34 (2): 169–184. https://doi.org/10.2307/26921125.
Menges, Leonhard. forthcoming. "Responsibility and Appropriate Blame: The No Difference View." *European Journal of Philosophy*. Advance online publication. https://doi.org/10.1111/ejop.12571.
Moore, Andrew. 2019. "Hedonism." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/hedonism/.
Nelkin, Dana K. 2016. "Accountability and Desert." *The Journal of Ethics* 20 (1): 173–189. https://doi.org/10.1007/s10892-016-9230-0.
Nida-Rümelin, Martine, and Donnchadh O Conaill. 2019. "Qualia: The Knowledge Argument." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2019. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/win2019/entries/qualia-knowledge/.

O'Connor, Timothy, and Christopher Franklin. 2018. "Free Will." In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Fall 2018. Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2018/entries/freewill/.

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

Pettit, Philip. 2007. "Responsibility Incorporated." *Ethics* 117 (2): 171–201.

Roskies, Adina L. 2012. "Don't Panic: Self-Authorship Without Obscure Metaphysics." *Philosophical Perspectives* 26 (1): 323–342. https://doi.org/10.1111/phpe.12016.

Schulte, Peter. 2014. "Beyond Verbal Disputes: The Compatibilism Debate Revisited." *Erkenntnis* 79 (3): 669–685.

Shoemaker, David. 2015. *Responsibility from the Margins*. New York: Oxford University Press.

Sripada, Chandra. 2016. "Self-Expression: A Deep Self Theory of Moral Responsibility." *Philosophical Studies* 173 (5): 1203–1232. https://doi.org/10.1007/s11098-015-0527-9.

Strawson, Galen. 1994. "The Impossibility of Moral Responsibility." *Philosophical Studies* 75 (1–2): 5–24.

Vargas, Manuel. 2013. "How to Solve the Problem of Free Will." In *The Philosophy of Free Will: Essential Readings from the Contemporary Debates*, edited by Paul Russell and Oisin Deery, 400–416. New York: Oxford University Press.

Wallace, R. Jay. 1994. *Responsibility and the Moral Sentiments*. Cambridge, MA: Harvard University Press.

Waller, Bruce N. 2011. *Against Moral Responsibility*. Cambridge, MA: MIT Press.

Watson, Gary. (1996) 2004. "Two Faces of Responsibility." In *Agency and Answerability: Selected Essays*, 260–288. New York: Oxford University Press.