Phenomenally Conscious Philosophical Zombies: How epiphenomenalism undermines free will

Wittenbols, Teun

Phenomenally Conscious Philosophical Zombies

How epiphenomenalism undermines free will

by

Teun Wittenbols

S3052966

Moral and Political Philosophy

Supervisor: Dr. V.A. Gijsbers

25 – 05 – 2024

Abstract

The debate on free will is long and expansive. The conventional debate revolves around the truth of determinism and whether it is compatible with free will. However, most of the discussions between compatibilists and incompatibilists seem to get bogged down on what it means to have freedom and end in a frustrating back-and-forth about semantics - a place which we so far have not been able to escape. In this paper, I approach this debate from a different perspective in order to avoid this familiar dead end. First, by drawing on our intuitions, I will argue that phenomenal consciousness is necessary for any conception of free will that is true to our intuitions. Next, I will explain the concept of epiphenomenalism: the view that phenomenal consciousness is an epiphenomenon and so has no causal power. Taken together, these ideas form an obvious problem for any free-willer: if phenomenal consciousness is essential for free will but it does not have any causal power, free will does not exist. It seems unreasonable to claim that adding phenomenal consciousness to a philosophical zombie (a being that functions exactly like a human being but has no phenomenal consciousness and so has no free will) makes him free. Anyone arguing for the existence of free will now faces a dilemma: they can either dispute the claim that consciousness is necessary for a conception of free will that is true to our intuitions, or they can dispute epiphenomenalism. The problem of disputing the claim that consciousness is a necessary condition for a conception of free will that is true to our intuitions is quite simple: I deem it impossible to give an example of an unconscious entity that has free will. Alternatively, disputing epiphenomenalism puts the burden of proof on the free-willers, since it would require denying the idea of *causal closure* in order to avoid the problem of *explanatory exclusion*. Defenders of free will would need to address *the hard problem of consciousness* and possibly even 'solve' it. The investigation in this paper therefore results in a robust conclusion about free will – that it is in fact an impossibility – and also provides a different perspective from which

Keywords: Free Will, Consciousness, Phenomenal Consciousness, Access-Consciousness, Epiphenomenalism, The Hard Problem, Causal Closure, Philosophical Zombies

Table of contents

1. Introduction

The debate on free will is long and expansive. Most of the canonical philosophers of the west have at some point in their works mentioned or discussed free will – Plato, Aristotle, Augustine, Aquinas, Descartes, Leibniz, Hobbes, Kant and Hume just to name a few (O'Connor & Franklin 2022). This is no wonder: few philosophical problems have as direct a connection to practical issues as this one. It underpins most – if not all – of our ideas about human action, psychology, morality, and punishment; and those form the basis for the way we structure our judicial and political systems, our educational and didactic institutions, and virtually all other aspects of society. Moreover, and perhaps more viscerally, it is deeply connected to our subjective experience of how we think and act and how we relate to and treat others within our own lives.

1.1 The Neuroscientific Revolution

The current increase in interest in the debate about free will might have to do with relatively recent developments in psychology and neuroscience. Findings from these fields from the last decades seem to indicate that our behaviour is much more explainable in terms of our genetic makeup, upbringing, and circumstantial factors than we might have thought (Sapolsky 2023). On the short-term neurological level, scientists such as Libet, Gleason, Wright & Pearl (1983), Haggard (2011), Haynes (2013), Fried, Mukamel & Kreiman (2011) seem to have been able to identify decisions or actions – or brain activity that necessarily leads to a decision or action – before participants stated they consciously decided to act. Although all these studies have received critique both on a scientific (Sanford et al. 2017; Fosu-Blankson et al. 2023) and philosophical level (Mele 2009, Dennett 2004), it is obvious why such findings make us question what freedom of the will would in fact look like and if it exists. On the long-term level we not only have a lot of neurological and psychological evidence on the predictability of almost all types of decisions and actions (Sapolsky 2023), the whole idea of doing research on what sort of influence someone's DNA, brain-structure, upbringing, external inputs, etc. has on their behavior presupposes that these elements do indeed influence behavior in a way that could be predicted to a certain degree.

Moreover, if we take a broader temporal perspective, we see that many of our societal institutions have adopted a view much more in line with some idea of explainable behavior. It is why we have 'mitigating circumstances' in our judicial systems and why we make use of the DSM-5: some facts about our brains and external factors directly or indirectly cause certain behavior. Although these findings do not at all prove that free will doesn't exist, they seem to point towards a view of behavior which is much less free than we might expect when thinking about free will. There seems to be some sort of disconnect between our – and with 'our' I do not mean philosophers, but I mean the general public – highly stylized conception of free will which I think is something along the lines of: 'humans can do whatever they want, and they can decide for themselves whatever it is that they want', and our way of describing human behavior. As soon as we learn that someone is schizophrenic, for example, we naturally view their behavior as less 'under their control' to a certain degree. If someone commits a crime, we often view them as less responsible if the circumstances that influenced their behavior were sufficiently dire.

1.2 Aim of this paper

Despite the scientific evidence, there are still steps to be made on a purely philosophical level. In this paper I will take a different perspective than is usual in arguing against the existence of free will. Instead of focusing on conventional terms such as *determinism*, *sourcehood*, and *the ability to do otherwise*, I will focus on *consciousness* and *epiphenomenalism*. Since, as Nadine Elzein puts it: ***"****In philosophy, while concerns about mental causation span back centuries, the question of whether epiphenomenalism undermines free will is surprisingly underexplored."* (Elzein 2019 p.3). Why is *the* important concept in the field of philosophy of mind – consciousness – so underexposed in the discussions on free will? In this paper, I will show that we can, in fact, use the concepts of *consciousness* and *epiphenomenalism* to draw robust conclusions about the existence of free will: it does not – and perhaps cannot – exist.

1.3 Structure of this paper

This paper will be structured as follows: first, in section 2, I will spell out the conventional debate on free will. I will set the stage by introducing relevant terms like *determinism* (2.1) and *quantum* (2.2) and the positions called *compatibilism* and *incompatibilism* (2.3). I will then explore multiple arguments for and against the existence of free will (2.4) and explain how the debate often gets bogged down on very specific concepts and conceptions of free will (2.5). This section's aim is to show that it is hard to say to what degree these are substantive or merely semantic discussions. Next, I will leave this conventional debate for what it is and approach the topic from a different angle. In section 3, I will highlight a specific element which plays an indispensable part in our intuitions about free will: *consciousness* (3.1). I will explain multiple conceptions of consciousness (3.2) and show which kind of consciousness is most applicable to the problem of free will (3.3). Then, in section 4, I introduce the philosophical position of *epiphenomenalism* – the view that consciousness is an epiphenomenon and has no causal power (4.1) – and give arguments for taking this position (4.2). In section 5, I highlight how combining these two ideas gives us reason to be highly suspicious of any account of free will (5.1) and puts the free will advocates for a dilemma: they can either deny the indispensability of consciousness for free will, or they can deny epiphenomenalism which puts the burden of proof on the free-willers. It would require denying the idea of *causal closure*, in order to avoid the problem of *explanatory exclusion*. It would need to address *the hard problem* and possibly even 'solve' it (5.2). Finally, I revisit some of the arguments from the conventional debate mentioned in section 2.4 and show that we can now view these arguments in light of our conclusions about consciousness and epiphenomenalism (5.3).

2. The Conventional Debate

Even though I specifically want to give an argument from a different perspective than is usual in the free will debate, I want to start by giving a brief overview of the conventional arguments and positions. This is useful to see why the current debate gets bogged down and why it is imperative to use different avenues to reach any conclusions about free will. Moreover, the argument I make in this paper does have specific consequences for people who hold specific positions in the conventional debate, and it is therefore useful to get familiar with them.

2.1 Determinism

Perhaps the most important concept in defining the possible positions in this debate has been the concept of *determinism*. Determinism, in essence, is the view that since the world is made up of stuff and that stuff behaves in ways that are explainable in terms of cause and effect and adhere to specific rules (the laws of physics) there is only one way the dice will eventually roll and one way in which the universe can play out. All that happens is determined to happen. A useful analogy was posited by Laplace back in 1814, which has come to be known as Laplace's demon. Sapolsky put the gist of the analogy concisely in his topical book *Determined* (2023):

> *"If you had a superhuman who knew the location of every particle in the universe at this moment, they'd be able to accurately predict every moment in the future. Moreover, if this superhuman (eventually termed "Laplace's demon") could re-create the exact location of every particle at any point in the past, it would lead to a present identical to our current one. The past and future of the universe are already determined."* (Sapolsky 2023 p.15)

Although we are not - and might never be - in a position to actually predict the future, the fact of the matter is (according to determinism) that only one future exists. Determinism already gives us some way of classifying philosophers: those who accept determinism and those who don't.

2.2 Quantum: an attack on determinism

An initial objection that is often raised questions determinism in light of our developing knowledge of quantum theory, because *"[…] the uncertainty and indeterminacy of the quantum world, according to the standard view of it, is not due to our limitations as knowers, but to the unusual nature of the physical particles themselves[.]"* (Kane 2001 p.7). Randomness on the quantum level suggests that there is not 'one future' we are heading towards, but many possible universes. In return, there are those who say that randomness itself cannot be a source of free will: if you let flipping a coin determine your next decision, you can hardly claim that your freedom in that decision resides in the coin-toss. If your next decision is the result of a chance quantum jump, how is this indicative of your free will? This leads free will deniers to state things like: "*Either our wills are determined by prior causes and we are not responsible for them, or they are the product of chance and we are not responsible for them."* (Harris 2012 p.5). Generally, I agree with the view that although quantum randomness does seem to undermine determinism, it does not so in a way that gives free-willers any tools. The only way in which I can see that it gives any ground for free will is when quantum jumps are not 'random' but are the result of something that *is* free will. There are philosophers who argue along these lines (Stapp 2009, 2011; Penrose & Hameroff 1996), however, this is beyond the scope of this paper so let's set quantum doubts aside and leave our idea of determinism intact.

2.3 Compatibilism

The more important question, however, is whether someone thinks free will and determinism to be compatible with each other or not. Those who think that they are compatible are aptly called *compatibilists*. Those who deem the two concepts to be incompatible (*incompatibilists*) can either dismiss determinism and accept free will (*libertarians*) or accept determinism and dismiss free will (*hard incompatibilists).* Of course, one could dismiss both determinism and free will which is a viable position, but those who do often have similar arguments as hard

incompatibilists along with claims about the irrelevance of quantum and chance as mentioned above.

At least on the surface, it seems quite logical to conclude that if determinism is true and we are moving somewhat helplessly towards a future which is set in stone, this conflicts with our ideas about free will: *the ability to do otherwise* is often stated as a necessary condition for free will (O'Connor & Franklin 2022), and determinism seems to implicate that we don't have the ability to do otherwise. This depends tremendously on what you mean by 'ability to do otherwise', of course. Another condition for free will which is often stated is *sourcehood*: you have free will when you are the true source of your actions (O'Connor & Franklin 2022). This term too has a plethora of different interpretations which we will discuss below. These elements – *determinism*, *the ability to do otherwise* and *sourcehood* – set the stage for the many philosophers who have contributed to the debate by proposing 'solutions' for the compatibilist position and those who have tried to undermine these 'solutions'. Let's look at a few of these arguments and where these discussions end up.

2.4 Where we get bogged down

I will now highlight a few major compatibilist solutions, incompatibilist counter arguments, widely used conditions for free will, and (very) specific definitions of terms. It is important to note that although I will try to convey these ideas clearly, my point is to show that these discussions often get very semantically technical and verge on incomprehensibility. We will leave these murky waters for the majority of this paper, but I first want to show that it is indeed necessary that we do. So, for now; sit tight.

A very common approach compatibilists take is arguing for free will by showing that people still have *the ability to do otherwise* in a determined world. G.E. Moore, for example, states that people have the ability to do otherwise in the sense that *"I could have walked a mile in twenty minutes this morning, but I certainly could not have run two miles in five minutes. I did not, in fact, do either of these two things; but it is pure nonsense to say that the mere fact that I did not, does away with the distinction between them, which I express by saying that the one was within my powers, whereas the other was not."* (Moore 1912 p.206). Similarly, he mentions that a cat could climb a tree, but a dog couldn't, regardless of whether they actually did or did not. According to Moore, the ability to do otherwise is best understood in a sense that has later been termed the *simple conditional analysis*:

> *"An agent S has the ability to do otherwise if and only if, were S to choose to do otherwise, then S would do otherwise."* (O'Connor & Franklin 2022)

Of course, incompatibilists are quick to point out that S could never *do* otherwise in any real sense because S cannot *choose* otherwise. The choices that result in our actions are themselves determined. In a suspicious move, the simple conditional analysis merely moves back the moment where freedom suddenly emerges – that is according to incompatibilists.

Vihvelin takes Moore's line of reasoning and builds upon it with the concept of finkish dispositions as conceived by David Lewis (1997). Free will, according to her, is the *"[…] ability to make choices on the basis of reasons and to have this ability is to have a bundle of simpler abilities."* (Vihvelin 2004 p.440) where she sees having an ability as *"having intrinsic properties*

that are the causal basis of the ability." (Vihvelin 2004 p.438). She states that the simple conditional analysis is flawed and revises it into the *revised conditional analysis:*

> *"S has the ability at time t to do X iff, for some intrinsic property or set of properties B that S has at t, for some time t' after t, if S chose (decided, intended, or tried) at t to do X, and S were to retain B until t', S's choosing (deciding, intending, or trying) to do X and S's having of B would jointly be an S-complete cause of S's doing X."* (Vihvelin 2004 p.438)

This would prove that free will exists according to Vihvelin. But of course, the incompatibilists would point to the fact that any ability we have, we have through no will of our own. So how would this make us free?

Another approach the compatibilists have taken – perhaps for the longest time – is to say that we have free will as long as actions are 'up to us' in a sense. As long as we are the source of our actions we have free will, regardless of whether this happens in a determined way. Arguments that take this line of reasoning have existed at least since debates between epicureans and stoics, where stoics posit that we have free will as long as actions come about 'through you' (O'Connor & Franklin 2022).

A very influential compatibilist argument was drawn up by Harry Frankfurt (1969) which seems to point to a way in which we can be said to be the source of our actions. In *Alternate Possibilities and Moral Responsibility* (1969) Frankfurt sketches a scenario where an agent 'Black' wants someone 'Jones' to choose a specific option of two. Jones can choose 'freely' between two things, if he chooses the one Black had in mind Black doesn't interfere, but as soon as Jones chooses the other option Black "[...] *takes effective steps to ensure that Jones decides to do, and that he does do, what he wants him to do."* (Frankfurt 1969 p.835). Others have come up with their own Frankfurt-style cases to make them more air-tight. Frankfurt had in mind a potion or some form of hypnosis, but Fischer (2006), for example, imagined that these '*effective steps'* could be conceptualized as a brain-chip implanted in Jones's head which can determine his decision and action. What is important is that Frankfurt claims that we feel Jones to be morally responsible for choosing the one option if he chose it himself but not if it was due to Black's interference. However, in this scenario we would hold Jones morally responsible if he chose the option himself even though he *could not have chosen otherwise*, because Black would've interfered. This seems to indicate that *the ability to do otherwise,* is not a sufficient condition for moral responsibility. The fact that Jones chose the one option on his own seems to make him the *source* of this action in a way that he would not be if Black had interfered, and "[w]*hen a person acts for reasons of his own* [...] *the question of whether he could have done something else instead is quite irrelevant*" (Frankfurt 2006 p.340). Even though this argument was originally only used to say something about *moral responsibility,* it is also narrowly related to certain conceptions of free will – this, again, depends on what we say free will is. *Sourcehood*, then, might be what free will is.

Incompatibilists rebut this by making reference to the so-called *consequence argument*:

> "*If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born [i.e., we do not have the ability to change the past], and neither is it up to us what the laws of*

> *nature are [i.e., we do not have the ability to break the laws of nature]. Therefore, the consequences of these things (including our present acts) are not up to us."* (Van Inwagen 1983 p.16, as cited in O'Connor & Franklin 2022).

Our intuitions about moral responsibility in the Frankfurt-style cases are themselves misguided since they rely on intuitions about free will, and that is exactly what the incompatibilists try to disprove. Jones has as much control over his choice if it was determined by Black's interference as when it was determined by other factors (Jones's brain, external factors, etc.) – the point is: all is determined and therefore outside of Jones's control. The compatibilists reply: that doesn't matter; Jones was the source of his action if Black did not interfere in a way he is not if Black did interfere.

One compatibilist approach to *sourcehood* is that of reasons-responsiveness: if an agent is sufficiently responsive to reasons that are available to the agent at the time of an action/decision he can be said to have free will (O'Connor & Franklin 2022). Fischer and Ravizza (1998) described the concept of *guidance control* which consists of "[…] *the idea that the mechanism that actually issues in the action must be the agent's own (in some sense to be specified), and the idea that this mechanism must be responsive to reasons (in a certain way)*" (Fischer & Ravizza 1998 p.62) where *"[t]he process by which an agent takes responsibility for the springs of his action makes them his own*[.]" (Fischer & Ravizza 1998 p.210). The incompatibilist will reply, again, that the agent has no control over whether they take responsibility for the springs of their action and no control over to what degree this mechanism is reasons-responsive.

Finally – at least for our present purposes – it is important to highlight the compatibilist idea of *sourcehood* as *identification*. This approach holds that as long as an agent is sufficiently identified with some of their motives or desires, they can be said to have free will. Many philosophers have come up with ways of conceptualizing what identification would precisely look like. Frankfurt (1971) introduces the term of *higher-order desires* and *higher-order volitions*: say you desire a piece of chocolate (a so called *first order desire*), but since you don't want to cheat on your diet, you want to *not* have this first order desire to eat chocolate. This 'want' about wanting is a *second order desire,* and when it becomes effective it becomes a *second order volition*. Frankfurt then gives an example of an addict who 'wants' to take drugs (*first-order desire*) but does *not* want to have this first-order desire (*second-order desire*). In this way, Frankfurt can stay that: "I*t is in virtue of this identification and withdrawal, accomplished through the formation of a second-order volition, that the unwilling addict may meaningfully make the analytically puzzling statements that the force moving him to take the drug is a force other than his own*[.]" (Frankfurt 1971 p.18). The addict is not identified with his first-order desire to take drugs and "[…] *it is not of his own free will but rather against his will that this force moves him to take it."* (Frankfurt 1971 p.18). When a person *does* identify with a desire (first- or higher-order) he has free will. The incompatibilist will, you guessed it, say that to what degree someone identifies with a desire is completely determined itself and so: is not free.

As we can now start to notice, the debate seems to get stuck on a rather definitional level. It is hard to keep track of what is actually meant by free will with philosophers such as Dennett arguing for a 'free will worth wanting', instead of just 'free will' (Dennett, 2004). There is debate

about determinism itself and what sort of consequences its truth would have. There are different conceptions of determinism and its relation to probability that seem to leave room for more than one possibility (Anscombe, 1971). There is discussion of what *abilities* are, and what *sourcehood* precisely means. There is much more to be said about all these topics and much has been said on these topics even in the last two decades, but I have become frightened that these discussions are destined to meet a dead end. This section must not be taken as an appropriate overview of the conventional debate, but as a way to get a feel for its inherent problems around semantics. I think it is clear now that it is imperative to escape this conventional debate.

2.5 Escaping the conventional debate

As we have seen, these various lines of argument lead to very complicated and technical positions. Much of the trouble, I think, is that free will is very hard to define in uncomplicated terms. All of the words which are used seem to have been defined from a free will perspective. For example, the whole concept of *choosing* depends on a certain form of freedom itself. *The ability to do otherwise* might in fact be a restatement of free will, but what is left of the definition of *ability* in a determined world? Compatibilists seem to give solutions where they simply state, 'this is what it means to have free will' and incompatibilists respond 'but how is that free?' – all in increasingly complicated terms.

Having gone through the trouble of dissecting these issues, it is clear that we need another way of discussing the problem of free will to make any progress. In the next section, I will characterize free will by highlighting a necessary and underappreciated component of what we call free will: consciousness. Then, once it is clear that consciousness is an indispensable part of what we call free will, I will present arguments for epiphenomenalism, the view that consciousness is an epiphenomenon and has no causal power. Combining these two ideas seems to make a very convincing case against the existence of free will.

So, let's forget about *determinism*, *sourcehood* and *the ability to do otherwise* for now and start our conceptual hike on a different path.

3. Consciousness and Free Will

> *"Many people may think they have a reasonably clear grasp of the concept of 'consciousness', for instance, or 'free will', but a little bit of probing usually reveals that providing a clear description or a set of defining characteristics is as difficult as providing a description of a specific colour, flavour or sound: we usually make do with a set of examples or comparisons*." (Slors, De Bruin & Strijbos 2015 p.10-11).

Since we are starting from scratch, we should start with giving a definition of free will. However, as we have seen in the last section, this step is problematic, and I want to avoid getting in semantic trouble. So, instead of giving a complete definition, I want to highlight one specific element which seems to play an important role in our intuitions about the concept: *consciousness*.

3.1 Consciousness: the key element

Even though we (clearly) don't all agree that some things *have* free will, I think we can all agree that some things *don't have* free will. Most of us don't believe passive objects have free will: a river (insofar as this can be seen as one 'thing') does not have any control over the way it flows. This might be so obvious that it may seem like a ridiculous starting point.

We humans are so different from rivers! For one, we are so much more complex: the way we have been shaped by evolution is incomparable to any passive object. We have nervous systems, digestive systems, muscles, and different senses, all working together to create one human being. Is complexity the required condition for some account of freedom? We are generally not inclined to say that mere complexity creates freedom. Let's take a computer: complex materials and intricate micro-infrastructure running millions of lines of code and hundreds of algorithms. Of course, humans may still be said to be much more complex. However, it seems to me that a computer does not have *any more* free will than a river. Complexity, moreover, is hard to define. A river, seen as a whole system, might be said to be *more* complex than a human being: having countless moving molecules all behaving according to the laws of fluid dynamics. A river is complex, but it has no free will. Complexity, then, is not the necessary element. What might it be that sets us human beings apart?

A river – however complex – is just a collection of water, and a computer – however complex – is a passive object, these are not living beings. Clearly, only living things could have free will. Okay, so we might shift our attention to a more complex and living 'being' shaped by evolution: a birch tree. Still, I suspect most people would not grant a tree free will. Can it decide in what direction it grows? No, it grows upwards, counter gravity, and towards the light. When looking at the detailed mechanics of how a birch tree grows, all is easily explicable in terms of a passive process where things 'just happen'.

Still, one might say, this is so far removed from us human beings: we perform actions and deliberate and think, we are agents whereas trees are not. So, what about animals? They perform actions. They deliberate and think, right? Here, I think we stray somewhat into a grey area regarding free will and I am less certain about people's intuitions. What is important to note here is that we generally think that animals are free in the sense that they can do what they want. This, however, is not yet free will. Frankfurt (1971), for example, only grants free will to beings with *second order volitions*: beings with the ability to effectively will a different will. The question then becomes whether animals have these *second order volitions*, which is hard to find out. Presumably, animals think and deliberate in different ways than humans do.

So, for clarity, let's leave animals aside for now and move to something as similar to humans as possible: humans. It is illuminating to see that we don't even grant humans free will in all cases. A sleepwalker does not have any control over where he walks and so we don't say that it is out of his free will that he walks a certain direction. A person who is hypnotized to perform a certain action, does not act out of his free will when performing that action. What is the relevant difference with 'normal' human behavior in these cases?

I want to draw your attention to a specific idea, which is that *consciousness* is the key element in our intuitions about free will. There seems to be no room for free will if 'the lights are off'. The

sleepwalker is not free in where he walks because it is not his consciousness but his subconscious at the wheel. The hypnotized person also seems to be acting on the basis of unconscious (or perhaps sub-conscious) processes. The specific characteristic of thinking and deliberating that intuitively sets humans apart from trees *is* – I think – consciousness. After all, what would unconscious thinking and deliberating look like? It seems a computer does exactly that, but we don't grant it free will. Why? Because it does so 'in the dark' – for all we know; a computer is not (self-)conscious.

This seems to me to clarify why animals are in a grey area; we just don't know exactly to what degree animals are (self-)conscious. The second order volitions Frankfurt (1971) mentions, for instance, seem only to work in conscious beings, since they have to be conscious of their first order desires, and we just don't know if animals are or can be conscious of those. Similarly, I think, we would place infants in this same grey area: do babies have free will? Only to the degree that we think they are (self-)conscious.

This is why studies like Libet's intuitively point towards the nonexistence of free will: since consciousness plays such a crucial role in our intuitions about free will, any study indicating that consciousness comes after the fact (Libet, Gleason, Wright & Pearl, 1983) endangers our conception of free will.

Our first premise, then, is that consciousness is necessary for a conception of free will that is true to our intuitions. It might not be a sufficient condition, but that is irrelevant for the rest of this paper. If we have grounds for believing that consciousness somehow doesn't exist, or that it *cannot* be the basis for free will – which is what I want to claim – we see that free will cannot exist.

3.2 What is consciousness?

In the last section, I have characterized free will by highlighting a specific and necessary element: *consciousness*. I have done so, instead of giving a full definition, to avoid getting into semantic problems – something which has troubled the conventional free will debate for a long time.

It is a little bit ironic that *consciousness* itself is one of the most elusive and strange concepts in the whole of philosophy. As David Chalmers stated in one of the seminal works in the field of philosophy of mind:

> *"Consciousness poses the most baffling problems in the science of the mind. There is nothing that we know more intimately than conscious experience, but there is nothing that is harder to explain."* (Chalmers, 2019 p.3)

However, I think that we don't need to get into too detailed descriptions of consciousness and address all the related problems (whether they be easy or hard). As we will see, we only need one specific conception of consciousness in order to draw conclusions about free will. I will here give a brief overview of some of the conceptions and explain which one is best applicable to the kind of consciousness necessary for free will.

There are various ways in which consciousness can be described, but a useful and influential distinction has been made between *phenomenal consciousness* and *access consciousness*

(Slors, De Bruin & Strijbos, 2015; Elzein 2019). This distinction was originally posited by Ned Block (1995).

The basic idea is that *phenomenal consciousness* has to do with the quality of subjective experience. In his seminal paper *What is it like to be a bat?* (1974) Thomas Nagel explained this best by highlighting the fact that what characterizes the consciousness of – say – a bat, is that *"there is something that it is like to be a bat."* (Nagel, 1974 p.438). The world appears to a bat in a certain way; there is a perspective there, where we would generally say that there is no such perspective in a passive object – like a baseball bat. There is not something it is like to be a baseball bat.[*] There are related conceptions of consciousness that we can classify under this header for now, like *qualitative* and *phenomenal states*, and *raw feels* (Elzein 2019).

Access-consciousness, on the other hand, has to do with the way in which certain things such as thoughts/idea/beliefs are *accessible* or *available* to you in such a manner that they can co-determine your actions, speech, and further thoughts (Slors, De Bruin & Strijbos, 2015). The way in which these things are conscious is better explained in the way they are accessible to you, than in terms of the way you experience them in terms of their quality.

It is important to note that these two ways of understanding consciousness are not mutually exclusive.

3.3 What kind of consciousness are we after?

Now, which way of understanding consciousness best fits the role it plays in our intuitions about free will? I think this highly depends on the specific sort of choice we are talking about. In the previously mentioned Libet study, a participant would be asked "[…] *when he felt like doing so, to perform the quick, abrupt flexion of the fingers and/or the wrist of his right hand*[.]" (Libet, Gleason, Wright, & Pearl, 1983 p.625). Some of the participants were even explicitly encouraged to be 'spontaneous'; to not actively deliberate about when to do it, but "[…] '*to let the urge to act appear on its own at any time without any preplanning or concentration on when to act'* [.]" (Libet, Gleason, Wright, & Pearl, 1983 p.625) Now, in such a stylized experiment, where you are basically waiting for a certain experience to arise, my feeling is that this is much better understood within the *phenomenal conscious* conception. It seems, at least intuitively, weird to say that the decision to flex your fingers is accessible to you and only at the moment of the decision gets accessed. It looks like this 'urge' you are waiting for is more like a *quale,* more like a feeling like 'pain' or 'itching'.

However, it is often objected that these sorts of inconsequential choices are not the most relevant cases for our discussions about free will: what matters are the choices we make in our lives where we really care about the outcome of the choice, like when considering what to study, whether to move to another country, or even on a more mundane level: what to eat, whether to go out for a jog or not, etc. These are the types of choices where we deliberate and

[*] *It is necessary to say that this is not an uncontroversial position to take. If one dives deeper into possible solutions to the hard problem – the problem of fitting phenomenal consciousness into a materialistic worldview – ideas like panpsychism and consciousness as a result of information processing seem to leave open the possibility of all objects and beings having this kind of consciousness.*

think, *consciously* – this seems more to be in line with the sort of consciousness that we saw was relevant for our intuitions about free will. We could conceptualize deliberating as thoughts and beliefs getting accessed as input for further thoughts and eventually those further thoughts are accessible to inform speech or co-determine action. In that sense deliberating is access conscious. However, here as well, the making of the choice itself cannot be conceptualized as something that is accessible. The making of a choice just happens as a result of a feeling of having deliberated enough and having an inclination towards one of the options.

So, although access-consciousness is relevant to free will, I take phenomenal consciousness to be the best candidate for the kind of consciousness which we highlighted was relevant for our intuitions about free will. We have the conscious experience of having some sort of control over our actions, we have the conscious experience of being able to do otherwise – these are not characterized by their accessibility for thought, speech, and action, these are characterized by their subjective quality.

This is further highlighted by the fact that we could think of an example of something which has access-consciousness without having phenomenal consciousness: a computer. A computer is able to access information which influences output, while I assume most people don't think that '*there is something that it is like to be*' a computer. Notice that I used this example previously to highlight that we intuitively say that a computer does not have free will, precisely because it does so 'in the dark' – it does so without phenomenal consciousness. I think this example makes clear that the kind of consciousness that is relevant for our intuitions about free will is *phenomenal consciousness*.

4. Epiphenomenalism

4.1 The concept of epiphenomenalism

> *"Epiphenomenalism is the view that mental events are caused by physical events in the brain, but have no effects upon any physical events."* (Robinson 2023).

In other words: epiphenomenalism is the idea that consciousness is an epiphenomenon. That implies that although consciousness is a truly existing 'thing' which emerges from neurons in the brain, it only emerges as a side product and has no causal force. Just as a shadow accompanies a traveler but doesn't influence their movements consciousness does so too (Jackson 1897). Or as Thomas Huxley put it:

> *"The consciousness of brutes would appear to be related to the mechanism of their body simply as a collateral product of its working, and to be as completely without any power of modifying that working as the steam-whistle which accompanies the work of a locomotive engine is without influence upon its machinery. Their volition, if they have any, is an emotion indicative of physical changes, not a cause of such changes."* (Huxley 1874 p.240)

This concept is thoroughly connected with one of the most groundbreaking ideas in the field of philosophy of mind: the *hard problem*. *The hard problem* is the problem of fitting the concept of

phenomenal consciousness into a physicalist worldview. This problem is larger than just epiphenomenalism and we don't need to address all the implications, but it is useful to view epiphenomenalism in light of this larger idea of *the hard problem*. David Chalmers coined the term in a chapter called *Facing Up to the Problem of Consciousness* of his book *The Character of Consciousness* (2010). He describes the problem as follows:

> *"Why is it that when our cognitive systems engage in visual and auditory information-processing, we have visual or auditory experience: the quality of deep blue, the sensation of middle C? How can we explain why there is something it is like to entertain a mental image, or to experience an emotion? It is widely agreed that experience arises from a physical basis, but we have no good explanation of why and how it so arises. Why should physical processing give rise to a rich inner life at all? It seems objectively unreasonable that it should, and yet it does."* (Chalmers 2010 p.5)

It is useful to bring up the idea of *causal closure* – or *causal completeness* – here. *Causal closure* is a theory which states that all physical events can be causally explained by other physical events. One neuron firing, for instance, can be fully explained by the firing of other neurons which have triggered it's firing. If causal closure is correct, and all physical events are caused by other physical events and nothing else, what is the function of phenomenal consciousness?

The idea is this best explained by making reference to a concept which arose in the field of philosophy of mind: the concept of a *philosophical zombie*. A philosophical zombie is an exact physical copy of a person. The only thing this copy lacks is conscious experience, it lacks *qualia* – in other words, it lacks phenomenal consciousness. The strength of this concept is not that it could really exist, but that it is at least conceivable. It would mean to have a human being who 'functions' exactly like other human beings but does so 'in the dark'. This begs the question: why would such a philosophical zombie not be possible? Why would phenomenal consciousness be necessary for us to 'function'?

Now, it is clear that consciousness *does* emerge: we have phenomenal consciousness – or at least I have.[*] For our purposes, the important point is the fact that we have this *hard problem* about phenomenal consciousness because there is no *further* reason for it to emerge: it has no function, it doesn't cause anything.

4.2 Arguments for epiphenomenalism

One of the strongest arguments for epiphenomenalism is the problem of *explanatory exclusion*. This idea comes from Jaegwon Kim (1989) but is best explained by Elzein:

> *"*[...] *if physics is "causally complete", every physical event has a sufficient causal explanation that appeals only to other physical events. This poses a problem: if we have a sufficient causal explanation of an event in physical terms, this seems to render mental phenomena causally redundant. Unless we are willing (implausibly) to posit constant overdetermination, the mental realm will turn out to be epiphenomenal*[.]*"* (Kim 1989, 1993, 1998, as cited in Elzein 2019 p.4)

[*] *The problem of solipsism is no hurdle for our discussions here. It would directly discount the possibility of free will for others and the question then remains: do I have free will? And since first-person phenomenal consciousness does emerge (at least for me, and if you are experiencing reading this: you too) this is the relevant scope for our discussions.*

What is meant by the 'mental realm' in this quote is *phenomenal consciousness*. Access-consciousness can work around this problem of explanatory exclusion by positing that mental states are *identical* to brain states. Mental states, then, are causally effective in the sense that the associated brain states are causally effective, which is totally compatible with the idea of causal closure (Davidson 1970). However, "[...] *even if this does make mental events causally effective, it does not make them causally effective qua their mental features*[.]" (Stoutland 1980; Honderich 1982; Sosa 1984, as cited in Elzein, 2019 p.4). The phenomenal aspect, the *what-it-is-like* aspect, is still unaccounted for. Philosophical zombies are perfectly conceptually possible in a causally closed physical world. They can even have access-consciousness. The real mystery is what role phenomenal consciousness plays in this causally closed world.

> *"If the physical world is, indeed, causally closed, as contemporary science has it, all physical events, including events in the pineal gland, have a complete physical explanation. The existence or inexistence of an immaterial soul would make no difference. Otherwise put: the soul would be out of a job when it comes to influencing our bodies and determining our behaviour."* (Slors, De Bruin & Strijbos, 2015 p.25).

If phenomenal consciousness has no causal power, it would make no difference whether it exists or not: it would be out of a job.

5. The Problem Epiphenomenalism poses for Free Will

5.1 The problem

Now we have concluded two things: phenomenal consciousness is necessary for any conception of free will that is true to our intuitions, and there is significant strength to the arguments supporting epiphenomenalism.

Let's return to our concept of the philosophical zombie. Does a philosophical zombie have free will? No. As can be seen from our first conclusion, phenomenal consciousness is necessary for any conception of free will that is true to our intuitions, and a philosophical zombie has no phenomenal consciousness – as a matter of definition. Like a sleepwalker – whom we don't grant free will – a philosophical zombie is moving without being phenomenally conscious of it. Like a computer – which we don't grant free will – a philosophical zombie has access to information and mental states that inform speech, thought, and action, but without any qualitative experience of them. A philosophical zombie 'goes through the motions', but does so in the dark.

Remember that in order to create our philosophical zombie, we have cloned a person and taken away the phenomenal consciousness of the clone. So, the only thing that sets these two beings apart is that only one has phenomenal consciousness. From the outside, we wouldn't be able to tell the philosophical zombie apart from the original person, since they would behave indistinguishably, and both seem conscious.

In order to properly compare these two beings – the person and philosophical zombie – we need to make sure they cannot interfere in each other's lives. So, while we're at it, let's imagine two identical universes, where we drop the original person in one of the two and the philosophical zombie in the other at the same exact moment and time. Let's assume further – only for the moment – that these two universes are subject to determinism. If we let these two beings live their lives in these two identical universes, they would lead the exact same life: they would make all the same decisions, at the exact same moments from the moment we drop them in these universes until the moment they perish at the exact same moment. At any point in their lives, if we were to take a snapshot of their universes, we would see two identical pictures. Whether the original person gets educated, starts a family, or buys a house, so will the philosophical zombie. Whatever sport the philosophical zombie plays, what movies they watch, and what music they will listen to, so will the original person.

Again, the only difference is that the original person is phenomenally conscious all the while, while the phenomenal zombie is not. We could say – in a cheeky oxymoron – that the original person is just a phenomenally conscious philosophical zombie.

It might start to become clear how the idea of epiphenomenalism thoroughly undermines our conception of free will. If the philosophical zombie does not have free will, how could it be that the phenomenally conscious philosophical zombie *does* have free will? If phenomenal consciousness does not influence any of the decisions and actions of the original person in any sense, how could it make him free?

To make the problem even more intuitively visceral, we could image a switch with which we can at any point in time turn on or off the phenomenal consciousness of the philosophical zombie. Imagine we turn on phenomenal consciousness every other day, while we turn it off in between. Now it has alternating days of being phenomenally conscious of every decision and action and being unconscious of the same things. Would we say this philosophical zombie has free will on the days where it is phenomenally conscious, and doesn't have free will on the days it isn't? This seems absurd.

All this seems to leave phenomenal consciousness a mere observant role. A one-person audience to the actions of the philosophical zombie. Like a movie the choices and actions of an entire life flash by in real-time to phenomenal consciousness without any way of interfering.

5.2 The dilemma for the free-willer

Now, anyone arguing for the existence of free will faces a dilemma: they can either dispute the claim that consciousness is necessary for a conception of free will that is true to our intuitions, or they can dispute epiphenomenalism.

The problem of disputing the first claim is that I think we have now established quite clearly how important the role is that consciousness plays in our intuitions about free will. I think examples like the sleepwalker and the computer illustrate this. The problem for someone who wants to argue for free will by disputing this claim is quite simple: give me an example of an unconscious entity that has free will. In fact, it doesn't even need to exist in reality: it could be as far-fetched as our philosophical zombie, but it has to be conceivable. I think this is quite impossible. It

would essentially be the same as granting a philosophical zombie free will, which I think proves that whatever remains of 'free will' in that case is not at all what we generally mean by it.

The problem of disputing the second claim is far more complex. This claim is so tied up with immensely complicated philosophical problems, that it puts the free-willer for multiple challenges. Disputing epiphenomenalism would require denying the idea of *causal closure*, in order to avoid the problem of *explanatory exclusion*. It would need to address *the hard problem* and possibly even 'solve' it, while as Chalmers stated: "[...] t*here is nothing that is harder to explain."* (Chalmers 2010 p.5). I think these challenges are so great that they might be insurmountable, but it at least puts the burden on the compatibilists and libertarians to show how phenomenal consciousness *could* influence the physical.

One way philosophers have tried to argue against epiphenomenalism has been to posit that mental states *are* brain states. This so-called *identity theory* has multiple interpretations but is best understood as saying that mental states are constituted by brain states (Slors, De Bruin & Strijbos 2015). As Ullin Place puts it: *"'Consciousness is a process in the brain', on my view is neither self-contradictory nor self-evident; it is a reasonable scientific hypothesis, in the way that the statement 'lightning is a motion of electric charges' is a reasonable scientific hypothesis."* (Place 1956 p.45). Philosophers taking this line of reasoning basically say that this hypothesis about consciousness just happens to be true (Slors, De Bruin & Strijbos 2015). This would mean that our philosophical zombies are in fact an impossibility since brain states (which our philosophical zombies have) *are* mental states. Epiphenomenalism would be untrue since mental states would have causal power *as* brain states.

However, there are some serious objections to *identity theory*. Most relevant for our discussion is the *explanatory gap*; similar to the theory of access-consciousness, identity theory seems unable to appropriately explain the subjective element of consciousness (Chalmers 2010; Levine 1983). As Levine noted, statements like "*Pain is the firing of C-fibers."* (Levine 1983 p.354) seem to explain how mental states like pain are constituted by physical events, but do not explain how pain feels. Identity theory, thus, lacks explanatory power for the subjective element of pain, which is exactly what phenomenal consciousness is. So, even if it would be true that philosophical zombies are in fact impossible, identity theory does not yet disprove epiphenomenalism since – again – "[...] *even if this does make mental events causally effective, it does not make them causally effective qua their mental features*[.]" (Stoutland 1980; Honderich 1982; Sosa 1984, as cited in Elzein, 2019 p.4). Identity theory does not show that *phenomenal* consciousness has any causal power.

5.3 Viewing the compatibilist in a different light

Now, before concluding, it is enlightening to revisit where we started: the confusing discussions between compatibilists and incompatibilists. I think our investigation and conclusion about consciousness and epiphenomenalism are quite useful in clarifying these discussions. We can now view compatibilist 'solutions' in a different light.

In the section on the conventional debate, we saw some compatibilist 'solutions' for free will and incompatibilist counterarguments to these proposed solutions. We can now see that any compatibilist solution to the problem of free will seems compatible not only with determinism,

but also with the concept of a philosophically zombie. *The ability to do otherwise* seems to apply to the philosophical zombie as much as to the phenomenally conscious philosophical zombie. If we are not willing to grant the philosophical zombie free will, then *the ability to do otherwise* seems to lose its power for explaining why the phenomenally conscious philosophical zombie can have free will. For example, if we revisit the *simple conditional analysis,* we see that a philosophical zombie would surely pass:

> *"An agent S has the ability to do otherwise if and only if, were S to choose to do otherwise, then S would do otherwise."* (O'Connor & Franklin 2022)

This goes for the *revised conditional analysis* too:

> *"S has the ability at time t to do X iff, for some intrinsic property or set of properties B that S has at t, for some time t' after t, if S chose (decided, intended, or tried) at t to do X, and S were to retain B until t', S's choosing (deciding, intending, or trying) to do X and S's having of B would jointly be an S-complete cause of S's doing X."* (Vihvelin 2004 p.438)

It seems that a philosophical zombie is perfectly able to meet both conditions while doing so completely 'in the dark', without there is *something it is like to be* that philosophical zombie. These cannot be conditions, then, for a conception of free will that is true to our intuitions. Our intuitions say: a philosophical zombie does not have free will.

Sourcehood, in a similar manner, can be ascribed to a philosophical zombie: an action can be 'up to' the philosophical zombie or 'come through them' just like it would be 'up to' the phenomenally conscious philosophical zombie, or 'come through them'. Let's return to the idea of reasons-responsiveness: if an agent is sufficiently responsive to reasons that are available to the agent at the time of an action/decision he can be said to have free will (O'Connor & Franklin 2022). What hinders our philosophical zombie to meet this condition? Acting on the basis of reasons seems to be a claim about having access-consciousness rather than phenomenal consciousness. Our philosophical zombie does have access-consciousness and can act on the basis of reasons, just like we might view a computer to act on 'reasons' or incentives. This too cannot explain why a philosophical zombie would not have free will, but a phenomenally conscious philosophical zombie – who would no longer be a zombie, of course – would.

Sourcehood as *identification*, on the other hand, is harder to explain in terms of consciousness. It seems that *identification* might in fact need an experiential element. It is hard to conceptualize *identification* in terms of access-consciousness, for example, since it seems to have to do with *feeling* identified with something. Identification is not something you merely have access to. This is not to say that reasoning along these lines is in any way promising to argue for the existence of free will, however, since even if identification – or rather: *feeling identified* – requires phenomenal consciousness, phenomenal consciousness still is an epiphenomenon and cannot cause anything. The phenomenally conscious element of identification, then, also cannot influence anything *qua mental feature*.

6. Conclusion

In this paper, I have argued against the existence of free will. The conventional debate around free will revolves around in what sense determinism is true and whether it is compatible with free will. As we have seen, most of the discussions between compatibilists and hard-incompatibilists seem to get bogged down on what it means to have freedom and end in a frustrating back-and-forth about semantics - a place which we so far have not been able to escape. I have approached this debate from a different perspective in order to avoid this familiar dead end. First, by drawing on our intuitions, I have argued that phenomenal consciousness is necessary for any conception of free will that is true to our intuitions. Next, I have explained the concept of epiphenomenalism: the view that phenomenal consciousness is an epiphenomenon and so, has causal power. Taken together, these ideas form an obvious problem for any free-willer: if phenomenal consciousness is essential for free will, but it does not have any causal power, free will does not exist. It seems unreasonable to claim that adding phenomenal consciousness to a philosophical zombie (who functions exactly like a human being but has no phenomenal consciousness and so: has no free will) makes him free. Anyone arguing for the existence of free will now faces a dilemma: they can either dispute the claim that consciousness is necessary for a conception of free will that is true to our intuitions, or they can dispute epiphenomenalism. The problem of disputing the claim that consciousness is a necessary condition for a conception of free will that is true to our intuitions is quite simple: I deem it impossible to give an example of an unconscious entity that has free will. Alternatively, disputing epiphenomenalism puts the burden of proof on the free-willers, since it would require denying the idea of *causal closure* in order to avoid the problem of *explanatory exclusion*. It would need to address *the hard problem* and possibly even 'solve' it. This investigation has not only resulted in a robust conclusion about free will – that it does in fact not exist – it also provides a different perspective from which to view the conventional debate. It puts compatibilist 'solutions' in a different light.

The conclusion that we have reached in this paper, that free will does not exist, has – or perhaps better put: *should have* – quite concrete consequences. One of the reasons that free will is such a hotly debated topic *is* the fact that, compared to some other philosophical debates, its falsity profoundly changes the way we should organize our societies and view our social interactions. For one, it should make us question whether concepts such as *blame* and *praise* are still valuable in our social interactions. If someone has no free will, can they still be held morally responsible for their actions? What basis is left for punishment? There are many philosophers who have discussed these issues (Harris 2012; Fischer & Ravizza 1998; Frankfurt 1969; Sapolsky 2023; Strawson 1962/2018) and although I will not go into their arguments here, it is important to note that what they all have in common is that their discussions are very much related to free will and whether it exists or not. All this to say: the conclusion that free will doesn't exist has great and far-reaching consequences that should make us rethink some of our most fundamental ideas about human interaction and society.

References

Anscombe, G. E. M. (1971). Causality and Determination. Cambridge University Press.

Block, N. (1995). On a confusion about a function of consciousness. *Behavioral And Brain Sciences*, *18*(2), 227–247. https://doi.org/10.1017/s0140525x00038188

Chalmers, D. J. (2010). *The Character of Consciousness*. https://doi.org/10.1093/acprof:oso/9780195311105.001.0001

Dennett, D. C. (2004). Freedom evolves. Penguin UK.

Elzein, N. (2019). Free Will & Empirical Arguments for Epiphenomenalism. In Virtues and Economics/Virtues and economics (pp. 3–20). https://doi.org/10.1007/978-3-030-26114-6_1

Fischer, J. M. (2006). My Way: Essays on Moral Responsibility (1st ed.). Oxford University Press, *Incorporated.*

Fischer, J. M., & Ravizza, M. (1998). Responsibility and Control. https://doi.org/10.1017/cbo9780511814594

Fosu-Blankson, F., & Inusah, H. (2023). A Critique of Libet and Wegner's Argument Against Free Will. *Zenodo (CERN European Organization For Nuclear Research)*. https://doi.org/10.5281/zenodo.774020

Frankfurt, H. G. (1969). Alternate Possibilities and Moral Responsibility. *The Journal Of Philosophy/The Journal Of Philosophy*, *66*(23), 829. https://doi.org/10.2307/2023833

Frankfurt, H. G. (1971). Freedom of the will and the concept of a person. The Journal of Philosophy, 68(1), 5. https://doi.org/10.2307/2024717

Frankfurt, H. G. (2006). Some Thoughts Concerning PAP. In D. Widerker and M. McKenna (eds.), *Moral Responsibility and Alternative Possibilities: Essays on the Importance of Alternative Possibilities* (pp. 339–445). Burlington, VT: Ashgate.

Fried, I., Mukamel, R., & Kreiman, G. (2011). Internally Generated Preactivation of Single Neurons in Human Medial Frontal Cortex Predicts Volition. Neuron, 69(3), 548–562. https://doi.org/10.1016/j.neuron.2010.11.045

Haggard, P. (2011). Decision Time for Free Will. Neuron, 69(3), 404–406. https://doi.org/10.1016/j.neuron.2011.01.028

Harris, S. (2012). Free will. Simon and Schuster.

Haynes, J. (2014). The Neural Code for Intentions in the Human Brain: Implications for Neurotechnology and Free Will. In The MIT Press eBooks (pp. 157–194). https://doi.org/10.7551/mitpress/9780262026680.003.0005

Honderich, Ted. 1982. The argument for anomalous monism. Analysis 42: 59–64.

Huxley, T. H. (2011). ON THE HYPOTHESIS THAT ANIMALS ARE AUTOMATA, AND ITS HISTORY [1874]. In Collected Essays (pp. 199–250). Cambridge University Press. https://doi.org/10.1017/CBO9781139149204.007

James, W. (1879). I.—ARE WE AUTOMATA? Mind, os-4(13), 1–22. https://doi.org/10.1093/mind/os-4.13.1

Kane, R. (2001). Free will. John Wiley & Sons.

Kim, J. (1989). Mechanism, purpose, and explanatory exclusion. Philosophical Perspectives 3: 77–108.

Kim, J. (1993). The non-reductivist's troubles with mental causation. In Mental Causation, ed. Heil & Mele. Oxford: Clarendon Press.

Kim, J. (1998). Mind in a Physical World. Cambridge, MA: MIT Press.

Laplace, P.S. (1951) A Philosophical Essay on Probabilities. French 6th Edition, Translation to English by F. W. Truscott and F. L Emory, Dover Publication, New York.

Levine, J. (1983). MATERIALISM AND QUALIA: THE EXPLANATORY GAP. *Pacific Philosophical Quarterly*, 64(4), 354–361. https://doi.org/10.1111/j.1468-0114.1983.tb00207.x

Lewis, D. (1997). Finkish Dispositions. *The Philosophical Quarterly (1950-)*, *47*(187), 143–158. http://www.jstor.org/stable/2956325

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). Brain, 106(3), 623–642. https://doi.org/10.1093/brain/106.3.623

Mele, A. R. (2009). Effective intentions. https://doi.org/10.1093/acprof:oso/9780195384260.001.0

Moore, G. E. (1912). *Ethics*, Oxford: Clarendon Press.

Nagel, T. (1974). What Is It Like to Be a Bat? *The Philosophical Review*, *83*(4), 435. https://doi.org/10.2307/2183914

O'Connor, T. & Franklin, C. (2022). Free Will. In E. N. Zalta & U. Nodelman (Ed.), The Stanford encyclopedia of philosophy (Winter 2021 ed.). Stanford University. https://plato.stanford.edu/entries/properties-emergent/

Penrose, R., & Hameroff, S. (1996). Conscious events as orchestrated space-time selections. *Journal* Of Consciousness Studies, 3(1), 36–53.

Place, U. T. (1956). IS CONSCIOUSNESS a BRAIN PROCESS? British Journal Of Psychology, 47(1), 44–50. https://doi.org/10.1111/j.2044-8295.1956.tb00560.x

Robinson, W. (2023). Epiphenomenalism. In E. N. Zalta & U. Nodelman (Ed.), The Stanford encyclopedia of philosophy (Summer 2023 ed.). Stanford University. https://plato.stanford.edu/entries/epiphenomenalism/

Sanford, P. D., Lawson, A. L., King, A. N., & Major, M. (2020). Libet's intention reports are invalid: A replication of Dominik et al. (2017). *Consciousness And Cognition*, *77*, 102836. https://doi.org/10.1016/j.concog.2019.102836

Sapolsky, R. M. (2023). Determined: A Science of Life without Free Will. Penguin.

Slors, M. V. P., De Bruin, L. C., & Strijbos, D. W. (2015). Philosophy of Mind, Brain and Behaviour.

Sosa, Ernest. 1984. Mind-body interaction and supervenient causation. Midwest Studies in Philosophy 9: 271–281.

Stapp, H. P. (2009). *Mind, Matter and Quantum Mechanics*. Springer Science & Business media.

Stapp, H. P. (2011). *Mindful Universe: Quantum Mechanics and the Participating Observer*. Springer.

Stoutland, Frederick. 1980. Oblique causation and reasons for action. Synthese 43: 351–367.

Strawson, P. (2018). Freedom and Resentment. In *Perspectives on Moral Responsibility* (pp. 45–66). Cornell University Press. https://doi.org/10.7591/9781501721564-002 (Original work published 1962)

Vihvelin, K. (2004). Free Will Demystified: A Dispositional Account. Philosophical Topics, 32(1/2), 427–450. http://www.jstor.org/stable/43154446

Wegner, D. M. (2002). The illusion of conscious will. MIT Press.