

RADBOUD UNIVERSITY NIJMEGEN



FACULTY OF SCIENCE

Is the Free Will Theorem about Free Will?

A PHILOSOPHICAL ANALYSIS OF THE FREE WILL THEOREM

THESIS BSc PHYSICS AND ASTRONOMY

Author:
Ella VAN DALEN

Supervisors:
dr. Luca CONSOLI
dr. Kian SALIMKHANI

Second reader:
prof. dr. Klaas LANDSMAN

January 2025

Contents

1	Introduction	3
2	Philosophical stances on free will and determinism	4
2.1	Determinism	4
2.2	Free will	5
2.3	Incompatibilism and the Consequence Argument	6
2.4	Compatibilism	8
2.5	The problem of free will and determinism	9
3	The Free Will Theorem	10
3.1	The axioms of the Free Will Theorem	11
3.2	The structure of the Free Will Theorem	12
3.3	Reception of the Free Will Theorem	13
4	Methods and approach	14
5	The ‘free will’ assumption	14
5.1	A free choice	14
5.2	Free will and the ‘free will assumption’	16
5.3	Philosophical free will and the undetermined choice	16
6	Free will for a human and a particle	17
6.1	The free will of a particle	18
6.2	Two interpretations of free will	19
7	Free will, the undetermined choice and indeterminism	20
7.1	The experimenter’s freedom	20
7.2	Variations of free will for a particle	21
7.3	Undetermined choice and randomness	22
7.4	Free will in an indeterministic world	23
8	The Free Will Theorem as an argument in the free will debate	24
8.1	The Free Will Theorem as an argument for libertarianism	25
8.2	Reformulation of the Free Will Theorem	26
8.3	Solutions to circularity	27
9	Conclusion	28

List of Figures

1	A table that shows the different positions in the free will debate.	9
2	A visual representation of the FWT and its relation to free will through the freedom of the experimenter.	16
3	A visual representation of the implication of the FWT in relation to free will. . . .	18
4	A visual representation of the circularity that should be avoided to be able to use the FWT as an argument against certain types of compatibilism.	26

1 Introduction

Whether free will exists is an extensive and much-discussed topic. This makes the debate on free will chaotic and makes for many possible miscommunications between disciplines. Although discussions about free will take place mostly within philosophy, science might be able to help out with some of the problems surrounding free will. Different disciplines of science could be helpful to the variety of questions in the debate about free will.

It seems that we make a free choice with our minds but how can this mental state then cause actions in the physical world? This is the problem of mental causation. To figure out how this could work, neuroscientists conduct experiments on the brain, because this could be the physical place for mental states. They try to see whether the brain might be the origin of our free will and control over our free actions. An important contribution to this is Libet's experiments [1]. For simple actions, like raising a hand, Libet showed in his experiments that before we have the conscious intention to act that way, an unconscious signal has already occurred in the brain. He concludes from his experiments that we are only free in the sense that we can choose whether or not to act out our unconscious impulses. Libet's experiments and conclusion have been debated controversially (e.g. by Gomes [2]) and it is doubtful that his interpretation is correct, but his experiments sparked many new theories some of which explain our acts as merely a result of the unconscious. For example, Wegner argues that free will is only an illusion [3]. He came up with experiments to show that humans can truly believe in reasons for why we act a certain way, even when these reasons are not true and we only came up with them afterwards. Therefore, we only believe that we made a free choice, while it is really just the unconsciousness that acted. So, experiments from neuroscience show evidence that our actions are maybe not as much up to us as we might like to think. But this does not yet give a definitive answer to the fundamental question of whether it is even possible to act in a way that is not completely determined by factors out of our control. As Libet says about the results of his experiments, 'But the deeper question still remains: Are freely voluntary acts subject to macro-deterministic laws or can they appear without such constraints, non-determined by natural laws and "truly free"?'.

So, the question remains, whether our free will can escape the determinism of the natural world. In the free will debate, there is a natural tension between free will and determinism. Since it seems that in a deterministic world, where everything including the outcome of your choices is determined, you could not have chosen to act otherwise. A famous attempt to prove this is the Consequence Argument by van Inwagen [4]. If this argument works, it is impossible to act freely in a deterministic world, as every choice would be determined already outside our control and you could not have done otherwise. However, it is still an open question whether this is the case. There are also various positions in the free will debate that try to find a way for free will to exist in a deterministic world. Overall, there are many different positions on whether free will, determinism or both are true. It is also not even sure whether the world is deterministic or indeterministic. This is a question about the fundamental nature of reality. Physicists do research into the fundamental laws of the natural world. So, theories of fundamental physics might also help in finding an answer to whether a deterministic world would allow for free will and whether the world is deterministic.

John Conway and Simon Kochen try to use physics and mathematics to prove something about free will and the determinism of the natural world in the Free Will Theorem (FWT) [5], [6]. They aim to prove that if humans have some type of freedom, then so do particles. It would be a great advancement if their proof could give direction in the philosophical debate on free will, but this does not follow directly from the result of the FWT. Their result, that particles are free, seems to hint at an indeterministic world. However, free will for the experimenter is assumed in the FWT. So, can this be a good argument for indeterminism and free will?

In this thesis, I will examine how the FWT can be used as an argument in the debate on free will and determinism. In the next chapter, I will first define free will and determinism in a way that is useful to this context. Then I will elaborate on the modern debate on free will and discuss some major positions in this debate. After this, I will give an overview of the proof in the FWT. This provides some of the information that is necessary to investigate how the FWT can fit into the debate on free will. To see how this would be possible, first I will look precisely at how the FWT is related to free will. I will also examine how freedom and determinism are related in the FWT. Then I will use this to interpret the FWT and see what possibilities and difficulties arise when trying to use the FWT as an argument in the free will debate.

2 Philosophical stances on free will and determinism

I will now present a brief overview of the philosophical debate on free will and determinism to later be able to interpret the FWT. In this, I will only focus on specific aspects of this debate, that are relevant to the further discussion in this thesis. First, a definition of determinism will be established, that is especially fruitful for discussing problems of free will. Then, with regard to the debate on free will and determinism, the meaning of free will will be defined. Finally, I will discuss some of the philosophical stances in this debate.

2.1 Determinism

In the debate on free will, the combination of free will and determinism is a controversial topic. To discuss this debate in the following sections, I will first look for a definition of determinism that is particularly fruitful in this debate. This is not a definition of determinism picked out for the FWT specifically, but one that is fruitful in the general discussion about free will. This then allows for the determinism of the FWT to be compared to the found version of determinism from the free will debate. So, a definition of determinism is needed, that makes the problem of free will and determinism exceptionally clear.

The concept of determinism has existed since ancient times and can be used in many different contexts. Therefore, many definitions of what determinism is exist. Very loosely the idea of determinism is that every event is necessitated by natural law that is applied to past events. Now to find a more precise definition of free will for our context, I will first discuss what we are looking for in a definition of determinism that is especially fruitful for the debate on free will. Andreas Hüttemann has proposed two criteria that such a definition of determinism should suffice [7]:

1. There is a challenge to free will due to determinism, that disappears in the case of indeterminism.
2. The source of this challenge to free will lies in the laws of nature or theories of the natural sciences.

Here indeterminism is the opposite of determinism. I will now go over different versions of determinism that are tested against Hüttemann's criteria.

A definition of determinism that has a chance of satisfying the two criteria is that of Jeremy Butterfield [8]. He wants to define determinism in a way that is related to physical theories, for the reason that physical theories are often a topic in the debate about determinism. He gets rid of causation and prediction in his definition of determinism because he argues that these are vague and controversial notions, that are not often used in physical theories. His definition depends on seeing

determinism as a property of theories. In physics, a theory usually describes a system that has properties, that are called states (a state in terms of the theory). The system also has regularities, that are called laws of the theory. Butterfield then states that a theory is deterministic, if and only if, for any two such systems, if they are in exactly the same state as one another at a given time, then according to the theory, they will be in the same state as one another at all future times.

Butterfield's definition of determinism cannot completely fulfil the conditions that were proposed by Hüttemann for a definition that is useful in the free will debate, because Butterfield's definition relates only to specific theories and not to the world. Two controversies arise when trying to expand Butterfield's definition of determinism to the entire world [8]. In the first place, the 'system' in question would be the entire world, which would require you to compare different worlds. It is not a given that this is possible. Secondly, a complete theory of the world needs to exist to expand this definition to the entire world. The possibility of such a theory is also not a given fact. Despite the problems Butterfield sees for applying theory determinism to the entire world, Earman did use a similar version of determinism for the entire world.

Earman defines determinism as a property of the world [9]. In his definition, a world is a four-dimensional space-time world. The actual world is the collection of all events that have ever happened, are now happening or will happen. A possible world is a collection of possible events that represent possible alternative histories of the actual world. W is the collection of all possible worlds that satisfy the natural laws of the actual world. Then for Earman, the world $w \in W$ is deterministic in case for any possible $w' \in W$ (so with the same natural laws as w), if w and w' agree at any time, then they agree for all times [9]. In this definition, if a world is deterministic, then natural law decides your actions, and therefore you cannot have free will. This definition fits Hüttemann's criteria. It suffices the first criterion because this determinism raises a challenge for free will that disappears for indeterminism, and the second one because the source of this challenge lies in natural law. So, according to Hüttemann, this is a fruitful definition of determinism to use for discussing free will. I will use this definition of determinism for discussing problems of free will in the upcoming chapters.

A world that is not deterministic is called indeterministic. Indeterminism means that natural law and initial conditions cannot uniquely determine the world at other times. It is not sure if we will ever know whether the world is indeterministic or deterministic. Determinism can be conceived of as a presupposition of science¹, it would then be an unnecessary pursuit to look for evidence of determinism, as you should just assume it. But determinism can also be seen as an empirical thesis, meaning it is an open question to science whether the world is deterministic. Hüttemann argues that determinism is most usefully conceived of as an empirical thesis [11]. Then it would in principle be possible for science to prove that the world is deterministic or indeterministic, but whether this can be done is another question. For the debate on free will, it is still an open question whether the world is deterministic.

2.2 Free will

There are a lot of strong beliefs about free will, as many people intuitively believe that they have at least some form of free will. Free will is also important to society, as it is often used to justify moral responsibility. After all, is someone guilty if they did not freely choose to act criminally? But without free will it would also not be possible to act differently to try to make a positive change in the world around you. Because free will is so important for humans, there are many different interests involved in the debate on free will, which makes the debate on free will even more complicated. But what is it that people mean when speaking of 'free will'?

¹See [10] for a discussion on whether determinism was seen as a presupposition of science historically.

I will first look at what Andreas Hüttemann means by free will because a definition of determinism was established based on his conditions about its relation to free will. Hüttemann takes the term ‘free will’ as an abbreviation of ‘the set of conditions that needs to be satisfied so that the agent in question has the control necessary for (moral) responsibility’ [7]. In this definition, he distinguishes between an epistemic condition and a control condition, that both need to be sufficed to have responsibility over your actions. While Hüttemann says that finding an exact definition of the epistemic condition is hard, to him this condition is based on the idea that an agent needs to have at least some knowledge about the results of their action to be (morally) responsible for this action. He defines the control condition as the condition that an agent should have at least some control over their actions. So, their actions should not be dictated by internal or external restraints that are out of their control, such as drugs or manipulation. He adds that historically an important part of what he calls the control condition is the Principle of Alternate Possibilities. The Principle of Alternate Possibilities states that a person is morally responsible for what they have done only if they could have done otherwise. Hüttemann’s definition of free will has conditions similar to the common definition in philosophy, but is ordered differently. In philosophy, it is common to split up the part of free will, that Hüttemann calls the control condition into two separate conditions. Then the condition that you could have done otherwise is a condition for free will on its own. The rest of the control condition can be stuck together with the epistemic condition into one other condition for free will. Despite the different ordering, Hüttemann’s demands for free will are the same as in philosophy. So, from now on the common definition of free will in philosophy can be used without raising problems for the way that determinism is defined.

It is widely agreed upon in philosophy that free will consists of the following two conditions [12].²

1. There must be different options to choose from. This is the condition that you ‘could have done otherwise’ (CHDO).
2. There must be an agent who chooses and acts out these choices.

The combination of these two conditions leads to problems regarding their relation to determinism [13]. The CHDO condition for free will seems impossible in a deterministic world.³ Since in a world where all our actions are determined by the past, it seems that you could not have done otherwise. But the condition of agency for free will does require some level of determinism⁴, because a choice without any reason or cause seems more random than free. Still, people do not want to give up their belief in determinism or free will. Whether and how determinism and free will can be combined is a central problem in the free will debate. This problem will be further discussed in the upcoming sections.

Concerning the FWT, it is also important to know what entities free will can be ascribed to. Most definitions of free will were formulated for humans (or possibly thinking animals). It is hard to pinpoint when exactly something can have agency and make a choice. Yet, it seems very hard to assign agency to a particle, as a particle does not have a thinking brain. So, a definition of free will that includes agency does not seem suitable for a particle.

2.3 Incompatibilism and the Consequence Argument

Incompatibilism is the idea that free will and determinism are not compatible. A famous argument for this is the Consequence Argument, which was proposed by Peter van Inwagen [4]. The argument

²Although, as will follow in section 2.4, some compatibilists do prefer to alter these conditions for free will.

³The Consequence Argument proves this, as in section 2.3

⁴This also depends on whether determinism is necessary for any sort of causation to be possible, which I will not go into here.

aims to prove that if determinism is true, a person could not have done otherwise. When seeing CHDO as a necessary condition for free will, the Consequence Argument then shows that free will and determinism are incompatible. For his argument, van Inwagen defines determinism in a similar way to the definition by Earman from section 2.1. He describes a world as deterministic when the state of the entire physical world at one instant, combined with laws that govern the world, uniquely entails all other states of the world.⁵

Van Inwagen [4] explains the Consequence Argument by the example of a judge who decides not to raise their hand at a time T_2 . T_1 is some time before the judge's birth and P_1 is the state of the world at T_1 . P_2 is the state of the world at time T_2 . L is the conjunction of all laws of physics. The argument then consists of seven statements (where statement (7) follows from the previous six premises):

1. If determinism is true, then the conjunction of P_1 and L entails P_2 .
2. If the judge had raised their hand at T_2 , then P_2 would have been false.
3. If (2) is true, then if the judge could have raised their hand at T_2 , they could have rendered P_2 false.
4. If the judge could have rendered P_2 false, and $(P_1 + L)$ entails P_2 , then the judge could have rendered $(P_1 + L)$ false.
5. If the judge could have rendered $(P_1 + L)$ false, then the judge could have rendered L false. (As the judge cannot render false states of the world before their birth.)
6. The judge could not have rendered L false.
7. If determinism is true, the judge could not have raised their hand at T_2 .

This argument shows that if determinism is true, you could not have done otherwise, if you cannot change the laws of nature or the past state of the universe. Thus, if free will is partly defined by the CHDO condition, free will and determinism cannot both exist. The idea that free will and determinism cannot both exist is called incompatibilism. There are different types of incompatibilism, as you can choose to believe in free will, determinism or neither. These will now be discussed shortly.

The position that neither free will nor determinism exists is called hard incompatibilism. Incompatibilists argue that free will does not make sense with or without determinism. However, this is not a very popular position, since the belief in at least free will or determinism seems favourable.

The position that the world is deterministic and free will doesn't exist is called hard determinism. In this case, no version of free will exists and the whole world is fully deterministic. However, an objection to this is that many people believe they have free will.

The position that the world is indeterministic and free will exists is called libertarianism. If free will exists in an indeterministic world, the question can be raised whether free will is freedom or merely randomness. It can then also be questioned whether agency is still possible for choice in an

⁵It is interesting to note here that van Inwagen himself uses 'laws of physics' in a broad way instead of 'laws of nature' for the laws that describe the world. He says he does this on purpose because he wants to exclude 'voluntary' laws about behaviour from determinism. By voluntary laws, he means psychological laws about the behaviour of rational agents, as he states in [4]: 'Let us, then, understand by "law of physics" a law of nature that is not about the voluntary behaviour of rational agents.' Yet, his state of the entire world, which is entailed by such physical law, does include the result of the act of a rational agent, like a man's hand that is being raised. So, van Inwagen's version of determinism will be interpreted as relating to everything in the physical world.

indeterministic world. An objection to free will in an indeterministic world is the ‘luck objection’, which argues that whether an undetermined choice occurs is just a matter of luck [14]. The choice then becomes just a random event and no agency can be included in this type of choice. Then indeterminism is not hospitable to a version of free will with agency. In this view, determinism might seem more hospitable to free will, since it does leave more room for agency. It is still a very open discussion whether the luck argument is valid.⁶

2.4 Compatibilism

The opposite to incompatibilism is compatibilism. Compatibilists believe that determinism and free will are compatible. There are different ways to approach this point of view, each version of which in some way has to dismantle the Consequence Argument. The Consequence Argument aims to prove that determinism excludes the possibility that you could have done otherwise. So a good start would be to see if the possibility that you could have done otherwise can be altered in a way such that the Consequence Argument is no longer a problem for free will. This is what classical compatibilists try to do.

For the Consequence Argument, a Categorical Analysis of CHDO was used. This analysis states that an agent only has the ability to choose otherwise than X at a time T, if and only if, holding everything up to T fixed, it was possible for the agent to choose or do otherwise than X. In this interpretation, CHDO is incompatible with determinism. Instead, classical compatibilists use a Conditional Analysis of CHDO [13]. This analysis states that you have the ability to do otherwise if and only if, were you to choose otherwise, then you would do otherwise. This changes CHDO to ‘I would have done otherwise, if I had wanted to’. This shifts the from acting in a certain way to wanting to act this way, of which the latter, they argue, is not a problem for determinism. Yet, the Conditional Analysis of CHDO is not seen as a sufficient argument for compatibilism. As it can be said that what you want also falls under determinism and you still could not have done otherwise. Additionally, this position has been refuted by sufficient counterexamples [17].

A new attempt to combine free will and determinism comes from new compatibilists. They do this by defining free will differently. They argue that CHDO is not a necessary condition for free will, so they put the emphasis solely on agency. An important contribution to this position is Strawson’s influential essay *Freedom and Resentment* [18]. In this essay, he argues that moral responsibility is fully based on reactive attitudes, which means our emotional responses to events. Then if free will is defined by moral responsibility, free will is not influenced at all by whether determinism is true or not. Another important contribution to this position is the paper *Alternate Possibilities and Moral Responsibility* [19] by Frankfurt, in which he argues that what is important for (moral) responsibility is the source of the actions, not whether someone really could have done otherwise. He shows with his Frankfurt cases that you can make a free choice to do something, even if the result of your choice would have happened anyway. Therefore, he argues that whether you choose to make a certain choice is more important than whether the outcome of this choice was already determined. So, both argue that responsibility for your actions is not determined by whether you could have done otherwise. New compatibilists then define whether a choice was made freely, by whether you are responsible for it. This version of free will, which is defined by responsibility (without the CHDO condition), then has no problem with determinism [12].

Contemporary compatibilists also attempt to combine free will and determinism, but they keep CHDO as a condition for free will in contrast to the new compatibilists. They try to find a way to dismantle the Consequence Argument by one of its premises, instead of directly altering CHDO like classical compatibilists do. To do this, they can either challenge the fact that you cannot change

⁶See for example [15] or [16] for discussions on this argument.

the past state of the universe (premise 5), or that you cannot change the laws of nature (premise 6). Many such theories have been proposed, of which a lot focus on counterfactual claims. These are claims that are not, but could have been.

A notable theory that allows that you could have done otherwise in a deterministic world is that of David Lewis [20]. He proposed a version of compatibilism, later called local miracle compatibilism by Helen Beebe [21]. Lewis argues that a person can act in a way such that, if they were to act in this way, some law of nature that does obtain, would not. He does this by distinguishing between the strong thesis and the weak thesis [20].

(Weak Thesis) I am able to do something such that, if I did it, a law would be broken.

(Strong Thesis) I am able to break a law.

Lewis argues that the weak thesis is true, while the strong thesis is not. So, he believes that a person cannot break a law, which would be performing a miracle. But he argues that they can act in a way that would require a miracle. He then undermines the Consequence Argument by arguing that premises 5 and 6 of the argument as in section 2.3 cannot both hold, as the weak or strong thesis would render one of them false. But it is still up to discussion whether he is successful in getting rid of the Consequence Argument. Beebe argues that the Consequence Argument is not severely damaged by Lewis' local miracle compatibilism and that he does not have proper reasoning to believe that the weak thesis is true, but the strong thesis is not [21].

2.5 The problem of free will and determinism

A definition of free will and determinism was found in the previous sections. There is a tension between the two, which results in many different positions on whether either one can exist. To summarize this debate, the major positions are shown in Figure 1.

		Free will	
		Yes	No
Determinism	Yes	Compatibilism	Hard determinism
	No	Libertarianism	Hard incompatibilism

Figure 1: A table that shows the different positions in the free will debate.

I will now quickly list the different positions in the free will debate. Incompatibilism encompasses the widely varying positions of libertarianism, hard determinism and hard incompatibilism. For incompatibilists, it is not a problem whether free will can exist in a deterministic world, because they believe in only one of the two. But a compatibilist tries to combine the two. A compatibilist argues that free will can exist in a deterministic world, so they have to explain how this would work. They run into the problem pointed out by the Consequence Argument, which is that it is not possible that you could have done otherwise in a deterministic world. Because free will is partly defined by the CHDO condition, this results in a problem for free will and determinism. This can be solved in various ways, which results in different theories within the position of compatibilism. There are new compatibilists (e.g. Frankfurt, Strawson) who argue that it is not necessary that you could have done otherwise for free will. It is also possible to alter what it means that you could have done otherwise to make this possible in a deterministic world. This is done in various ways, of which a notable version is Lewis' local miracle compatibilism.

It is important to highlight an assumption that underlies some of the discussed problems that occur for free will in a deterministic world. It is assumed that in a deterministic world, also humans are determined by natural law. This follows from, as Hüttemann puts it [7], 'the assumption – and in this respect I follow Spinoza – that the behaviour of all systems, if it is determined at all, is determined by fundamental laws of nature.' He calls this view Spinozaic naturalism. Without this assumption, if human behaviour falls outside natural law, human behaviour would also not be limited by natural law. The contradiction between human free will and determinism then lifts, because human behaviour is no longer governed by the determinism of the natural world. While this might seem like a good solution for the free will debate, it would require humans to exist or act at least for some part outside of their physical bodies. Such a position is taken by Descartes, who argues for mind-body dualism, which means humans have an immaterial soul in addition to their material bodies [12]. Then the mind is not a part of the natural world and can be free despite the determinism of the natural world. In this view, having freedom outside the deterministic natural world is limited only to humans. A big problem with this view is the lack of explanation of how an immaterial mind could influence the actions of a body in the material world. This is one of the reasons that mind-body dualism is not a popular belief in the modern debate on free will [12]. It also seems that if the freedom of the mind can cause changes in the material body that is part of the natural world, it raises a problem for determinism in the natural world. Since the mind would then allow for the possibility of undetermined change in the supposedly deterministic world. This shows why Spinozaic naturalism is part of what makes it possible for the entire world to be deterministic in a way that causes a problem for free will. In the view of Spinozaic naturalism, a challenge to free will occurs in a deterministic world, because if the human body is part of the natural world under determinism, every last change in the human body should also be determined. So, discuss the FWT and the free will debate, in this thesis we look at the problem of free will in a deterministic world (as defined by Earman) under Spinozaic naturalism, where the world, including the behaviour of humans, is determined by the laws of nature.

3 The Free Will Theorem

John Conway and Simon Kochen first published the Free Will Theorem in *Foundation of Physics* [5] in 2006 and they published a stronger version of the theorem in *Notices of the American Mathematical Society* [6] in 2009. In the FWT Conway and Kochen aim to prove that if humans have free will, elementary particles also have a small share of this. The original proof of the theorem is based on three axioms, SPIN, TWIN and FIN, of which SPIN and TWIN are based on quantum mechanics

and FIN is taken from relativity theory. In the strengthened version from 2009, the FIN axiom is replaced by the MIN axiom. Since the strong FWT surpasses the older version, the strong version will be examined. To set the stage for further discussion of the FWT in the upcoming chapters, I will first go over the axioms in the FWT and how its proof is constructed.

3.1 The axioms of the Free Will Theorem

The FWT is based on a thought experiment, where two space-like separated experimenters perform a squared spin measurement on a spin 1 particle. I will now discuss the settings of this experiment, by going over the three axioms, SPIN, TWIN and MIN, and the Kochen-Specker theorem, which all together make the FWT possible.

The first axiom is the SPIN axiom, which tells something about the results of performing a squared spin measurement on a spin 1 particle.

The SPIN Axiom: Measurements of the squared (components of) spin of a spin 1 particle in three orthogonal directions always give the answers 1, 0, 1 in some order.

The SPIN axiom follows from the quantum mechanical prediction that the squared spin operators in three orthogonal directions for a spin 1 particle commute and have sum 2. Thus, the squared (components of) spin of a spin 1 particle in three orthogonal directions always give the answers 1, 0, 1 in some order. This also entails two additional properties for three orthogonal spin measurements. The first property is that opposite directions have the same value and the second is that two orthogonal directions are never both 0.

From the SPIN axiom follows another part of the proof of the FWT, a theorem that Conway and Kochen refer to as the ‘Kochen-Specker paradox’. This theorem was proposed by Kochen and Specker [22]. It shows that the squared spin component of a spin 1 particle in a certain direction does not exist prior to measurement. Suppose the value of the measurements exists prior to measurement, then there would be some function for the outcome of the measurement on the unit sphere (that contains all possible directions for measurements). This function always has to suffice the properties of the outcome of a measurement in three orthogonal directions, that were a consequence of the SPIN axiom. A function on a set of directions that has all three of these properties is called a ‘101-function’. The Kochen-Specker theorem proves that no such function can exist for Peres’ 33-direction configuration. Peres’ directions are the directions that are used in the Kochen-Specker theorem, but this also suffices for the proof of the FWT. From these 33 directions follow the 40 triples of directions that the experimenter in the FWT can choose from. I will not go into the technical details of the Kochen-Specker theorem, as these are not very important to the discussion in the following chapter.

The second axiom is the TWIN axiom, this also follows from quantum mechanics. The TWIN axiom is based on a property of quantum mechanics that Einstein, Podolsky and Rosen first pointed out in their paper on what is now called the EPR paradox [23]. They say that it is possible to produce a pair of spin 1 particles, that both give the same answer to squared spin measurements in parallel directions. Particles in such a pair are called ‘twinned particles’. Twinned particles will give the same result to the same measurement, even if no information can pass between them between both measurements, as is the case in the FWT. In the FWT no information can pass between particles because they are space-like separated and information cannot travel faster than the speed of light. The TWIN axiom describes that the particles in the FWT are twinned.

The TWIN Axiom: For twinned spin 1 particles, suppose experimenter A performs a triple experiment of measuring the squared spin component of particle a in three orthogonal directions x , y , z , while experimenter B measures the twinned particle b in one

direction, w . Then if w happens to be in the same direction as one of x, y, z , experimenter B's measurement will necessarily yield the same answer as the corresponding measurement by A.

The third and final axiom is the MIN axiom, that relates to the measurement settings and their influence on the experiment. In contrast to the latter two axioms that follow from quantum mechanics, this axiom follows partly from relativity theory. Part of this axiom is a result of the paradox in relativity, that temporal order depends on the choice of inertial frame. This is combined with the principle that events from the future cannot influence the past. From putting this together, it follows that the choice of measurement directions of one experimenter cannot influence the response of the particle on the side of the other experimenter. The second part of this axiom is the assumption that experimenters can freely choose between experiments, where free choice means a choice that is not a function of what has happened in the past.⁷

The MIN Axiom: Assume that the experiments performed by A and B are space-like separated. Then experimenter B can freely choose any one of the 33 particular directions w , and a's response is independent of this choice. Similarly and independently, A can freely choose any one of the 40 triples x, y, z , and b's response is independent of that choice.

The FWT then follows from the described thought experiment, where two space-like separated experimenters perform a squared spin measurement on a spin 1 particle. It then turns out, under the three described axioms, that if the measurement of the particle were a function of the past, it would have to be a 101-function.⁸ This 'functional hypothesis' is rejected by the Kochen-Specker theorem. Thus, it is concluded that the particle's response is not a function of what has happened in the past.

The Free Will Theorem. The axioms SPIN, TWIN and MIN imply that the response of a spin 1 particle to a triple experiment is free—that is to say, is not a function of properties of that part of the universe that is earlier than this response with respect to any given inertial frame.

3.2 The structure of the Free Will Theorem

Now I will examine the steps in the proof of the Free Will Theorem to see where controversy could arise. The argument can be deconstructed in the following steps, where if the axioms in steps (1) to (3) are true, it follows from the Kochen-Specker theorem (4) that the result (5) is also true.

1. SPIN axiom
2. TWIN axiom
3. MIN axiom
 - a. An experimenter's choice is not a function of the past.
 - b. Independence of a particle's result of the measurement settings of the other particle.

⁷'Past' universe is used as meaning 'that part of the universe that is earlier than this response with respect to any given inertial frame'.

⁸The mathematical details of this proof I will not go into, as they are lengthy to describe and not very relevant to the discussion in the upcoming chapters.

4. Kochen-Specker theorem
5. The particle's response is not a function of the past.

In step (3) the MIN axiom is split up.⁹ The assumption that an experimenter's choice is not a function of the past in (3a) is noted separately from the assumption of independence of measurement result of one particle from the settings of the other particle in (3b). The assumption (3a) that the experimenter's choice is not a function of the past, will be referred to as the 'free will assumption'.¹⁰

The SPIN axiom (1) and the TWIN axiom (2) follow from a small corner of quantum mechanics. Part (3b) of the MIN axiom follows also from physics but from special relativity instead of quantum mechanics. As Conway and Kochen themselves argue, and Landsman also notes [25], very few physicists would reject these assumptions. The Kochen-Specker theorem is also a well-accepted and proved theorem. As the FWT is an attempt to prove something philosophical from physics and mathematics, it is also logical to accept the axioms that follow from some of the strongest theories of modern physics as a starting point. This leaves only part (3a) of the MIN axiom to be questioned, as this is not a direct result of physics. So, granted the axioms based on modern physics, belief in the FWT depends on belief in assumption (3a) that an experimenter's choice is not a function of the past.

For the following chapters, the free will assumption (3a) will be of most interest. First of all, it can be questioned whether this is a reasonable assumption. Conway and Kochen mention that they treat the ability of an experimenter to make a choice that is not a function of the past as a tacit assumption [6]. They give no explicit reason why this assumption should be accepted. Secondly, it can be questioned what exactly is meant by the free will assumption. As Conway and Kochen relate the FWT to free will through the name of this assumption but their formulation of this assumption, as the ability to make a choice that is not a function of the past, might not be the same as free will. Finally, it can be asked why this specific assumption is part of the FWT at all. Possibly, the combination of the axioms that follow from physics with another assumption, could lead to very different results. This could limit the impact of the FWT.

3.3 Reception of the Free Will Theorem

I will now give a brief overview of critiques that the FWT has received. The FWT has been criticized in multiple ways within the physics community for its impact on the interpretation of quantum mechanics. Various papers conclude that the FWT does not succeed in showing relativistic GRW-type theories impossible [26],[27],[28],[24]. In addition to that it is argued in [28] that the FWT lacks novelty compared to Bell's inequality [29]. It is even argued in [28] that from the same experiment similar results as in the FWT have been found before in [30], [31], [32], [33] and [34]. The novelty of the FWT that is of interest now is the focus that the FWT has on free will, it tries to take the result further than just physics. The focus of this thesis will be the philosophical aspects of the FWT, specifically its impact on the free will debate. So, what has been said about this?

Wüthrich argues that the FWT does not provide sufficient proof for indeterminism [24]. Which limits how the FWT can impact the free will debate. Also, the different comments by Walleczek [35] and 't Hooft [36] on the type of free will in the free will in the FWT, can be limiting for the impact of the FWT. One way to use the FWT as an argument in the free will debate has been proposed. After a reformulation of the FWT by Cator and Landsman [37], Landsman shows that the FWT can be used as an argument against local miracle compatibilism [25]. This result will be

⁹This split agrees with the distinction Wüthrich makes between parts of the MIN axiom in [24].

¹⁰This is the same as how Conway and Kochen call this assumption in their original version of the FWT.

further discussed in section 8.3. Apart from this no other attempt to use the free will theorem as an argument for a position in the free will theorem has been found, apart from Conway and Kochen's claims, which can be regarded as controversial. In this thesis, I will give an additional conceptual analysis of some of the philosophical claims made in Conway and Kochen's papers on the FWT.

4 Methods and approach

In the FWT Conway and Kochen aim to prove that if a human can choose freely, a particle can also respond freely. Although this in itself does not show any position in the free will debate, there might be possibilities for the FWT to steer the debate on free will in the direction of indeterminism. In this thesis, I will investigate whether and how it is possible to use the FWT as an argument in the free will debate.

To see how the FWT can be used as an argument in the debate on free will, I will do an argumentative analysis of the FWT. For this, I will look at the philosophical claims that Conway and Kochen make in their papers on the FWT. I will look precisely at the definition of terms like 'free will' or 'determinism' that they use for this. I will see whether they are consistent in their use of these terms. Then I will compare this to definitions in the philosophical debate, so I can get insight into how the FWT fits into this debate. This way I can see if their understanding of concepts like 'free will' might result in problems for the arguments they make. While doing this I will also keep in mind any assumptions made in the FWT and see how these restrict its impact. From all this, possible problems to use the FWT as an argument in the free will debate will come to light. Then I will look for solutions to circumvent these problems.

In the following chapters, the described analysis of the FWT will be executed. I will first discuss in Chapter 5 what type of freedom the free will assumption consists of and how this assumption relates to free will. Then in Chapter 6, I will see how Conway and Kochen view the freedom of the experimenter and the particle, and how this combines with the philosophical definition of free will. From this, different interpretations of free will arise. In Chapter 7 I will see how Conway and Kochen interpret free will in the FWT, and what reasons they might have for this. Finally, in Chapter 8 I will use said definitions and interpretations to see how the FWT can be used as an argument in the free will debate. I will examine the problems that arise and discuss possible solutions.

5 The 'free will' assumption

The type of freedom that is relevant to the FWT is that of the free will assumption, which states that an experimenter has the ability to make a choice that is not a function of the past (assumption (3a) from section 3.2). This is the type of freedom that is assumed for the experimenter and that is aimed to be proved for a particle. In this section, I will look at how Conway and Kochen define the free choice, that the experimenter can make according to free will assumption. I will discuss what it means to be able to make a free choice. Then I will compare the free will assumption to free will as defined in philosophy.

5.1 A free choice

The free will assumption states that the experimenter has the ability to choose their measurement settings freely. This is explained more specifically by Conway and Kochen as follows [6] (p. 228).

To say that A's choice of x , y , z is free means more precisely that it is not determined by (i.e., is not a function of) what has happened at earlier times (in any inertial frame).

Conway and Kochen define what it means to be free for a choice or response as not being determined by what has happened at earlier times (in any inertial frame).¹¹ How and why do they define being free as being undetermined by what has happened in the past?

In the FWT ‘not being determined by’ is equivalent to ‘not being a function of’. The formulation of ‘not being a function of’ follows directly from the proof of the FWT, as this proof depends on a contradiction in the ‘functional hypothesis’. The functional hypothesis describes the case in which the particle’s response would be a function of the past. This response should then be a 101-function of information¹² from earlier than the response. The Kochen-Specker theorem shows that this specific type of function cannot exist, thus making the functional hypothesis impossible. This is then what proves that the response of the particle is not a function of what has happened in the past. The use of the phrase ‘is not a function of’ is then easily explained, because it is directly related to what the theorem proves. ‘Not a function of’ can be replaced by the equivalent ‘undetermined by’,¹³ which sounds possibly a little less mathematical. To relate the theorem to freedom, a choice that is undetermined by past events is called a ‘free choice’ by Conway and Kochen.¹⁴

The free will assumption (assumption (3a) in section 3.2) is not just about a free choice in itself, in this assumption the experimenter who makes the choice is also incorporated. So, a free choice is a choice that is undetermined by what has happened in the past, but how does the free choice relate to the person who makes it?

In their definition of what it means for a choice to be free, Conway and Kochen say that there is an agent involved who makes the free choice, by stating in their definition of a free choice, that ‘[experimenter] A’s choice’ is free. The experimenter then can choose freely in the sense that they can make a free choice as Conway and Kochen defined this. So, what it means to be able to ‘choose freely’ is that you have the ability to make a choice that is undetermined by what has happened in the past. It is in this case not the agential control that makes the choice free. So, agential control is not a necessary part of what makes this choice free.¹⁵ The freedom that the FWT assumes for the experimenter, consists only of their ability to make a choice that is undetermined by what has happened in the past.

The FWT aims to prove that if an experimenter can make a choice that is free in this sense, then a particle’s response is free in the same sense. They describe this in the paper on the strong FWT as follows [6] (p. 228).

Our theorem is the surprising consequence that particle a’s response must be free in exactly the same sense, that it is not a function of what has happened earlier (with respect to any inertial frame).

When regarding the particle, Conway and Kochen talk not about a choice, but only about ‘[particle] a’s response’. In this case, the response of the particle is what is free, in the same sense the choice of the experimenter is free, which is not being determined by the past. Again, like the experimenter, the particle is free only in the sense that it has the ability to respond in a way that is undetermined by what has happened in the past.

So, the freedom that the experimenter and the particle have in the FWT is solely the ability to make a choice that is undetermined by what has happened in the past. With the free will assumption

¹¹In the rest of this thesis ‘what has happened in the past’/‘past events’ will be used equivalently to ‘what has happened at earlier times with respect to any inertial frame’.

¹²‘Properties’, ‘events’ and ‘information’ are used interchangeably by Conway and Kochen in the FWT.

¹³Whether the phrase ‘undetermined by’ relates to some version of determinism will be discussed in section 7.4. This question will become important for the philosophical impact of the FWT.

¹⁴Whether an undetermined choice is really free is discussed in section 7.3.

¹⁵The possible reasoning and consequences for such a definition of freedom are discussed in Chapter 7.

this freedom is assumed for the experimenter and then the FWT aims to prove that the particle has this same type of freedom. The reasoning and impact of such a definition of freedom for the experimenter and particle is further discussed in Chapter 7.

5.2 Free will and the ‘free will assumption’

To see in what way the FWT is about free will, I will examine how Conway and Kochen relate the free will assumption to free will. So, I will first discuss how they relate the freedom of the experimenter to free will.¹⁶ Conway and Kochen claim that their theorem proves that if humans have free will, elementary particles also have some share of this in the strong FWT [6] (p. 226).

It asserts, roughly, that if indeed we humans have free will, then elementary particles already have their own small share of this valuable commodity. More precisely, if the experimenter can freely choose the directions in which to orient his apparatus in a certain measurement, then the particle’s response (to be pedantic—the universe’s response near the particle) is not determined by the entire previous history of the universe.

So, they say that for humans to have free will means more precisely that the experimenter can choose freely. In other words, they say that if humans have free will, the experimenter has the ability to make a choice that is undetermined by the past. So, they replace their free will assumption that relates only to the ability to make an undetermined choice, by the assumption that an experimenter has free will in a much broader sense. This shows that Conway and Kochen make the implicit assumption that if a person has free will, this person can make a choice that is undetermined by what has happened in the past. This is how they claim to relate the FWT to free will, which is shown in Figure 2.

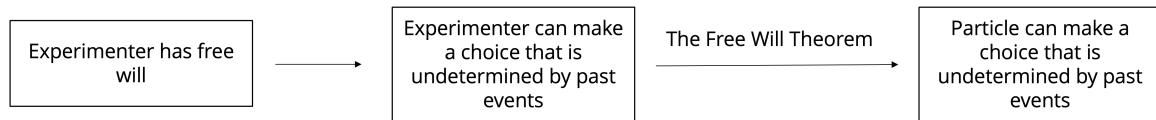


Figure 2: A visual representation of the FWT and its relation to free will through the freedom of the experimenter.

5.3 Philosophical free will and the undetermined choice

To be able to see how the FWT can be used in the debate on free will later, it is now worth discussing whether the meaning of ‘free will’ as how it is used by Conway and Kochen is the same as the free will that is used in philosophy. I will now discuss whether philosophy would agree with Conway and Kochen’s use of free will as described in the previous section. So, we will look at whether it is reasonable to say, from a philosophical point of view, that the ability to make an undetermined choice follows from free will. If this were the case, the ability to make an undetermined choice should first of all not in itself undermine free will, and secondly be a necessary consequence of free will. So, how does the undetermined choice combine with the conditions of free will in philosophy?

In philosophy, the two conditions for free will are the conditions of CHDO and agency (as discussed in section 2.2). The freedom of the experimenter in the FWT is defined by their ability to make a choice that is not determined by what has happened in the past, not their agency. The ability to make an undetermined choice that Conway and Kochen describe conflicts with agency for

¹⁶Free will for the particle in the FWT will be discussed in Chapter 6.

those who believe that agency is excluded by an undetermined choice, but this is still an open debate as discussed in section 2.3. For now, it is granted that the free choice of the experimenter, although not free because of agency, could still have been made under agential control. Now for the second condition of free will, if a choice is undetermined, this allows that you could have done otherwise. So, the ability to make an undetermined choice does not in itself undermine free will, but is it also a necessary condition for free will? In other words, does the ability to make an undetermined choice follow from either the agency or CHDO condition?

The experimenter in the FWT could have done otherwise since the outcome of the choice was not determined. The other way around, a free-willed person, who could have done otherwise, has different options to choose from. So, their choice is not determined by what has happened in the past. Then, if CHDO is a necessary condition for free will, it also follows from your free will that you have the ability to make a choice that is not determined by what has happened in the past.

Yet, CHDO for free will is not accepted by all philosophers as a necessary condition for free will. As, according to the Consequence Argument, this condition results in a problem with determinism. Therefore, new compatibilists (like Strawson and Frankfurt) argue that CHDO is not a necessary condition for free will. So, for them, free will does not enable you to make a choice that is not determined by what has happened in the past. From this point of view, the FWT is not at all related to free will. Other compatibilists accept CHDO as a condition for free will. They argue that you have options to choose from, even in a deterministic world. But for them, that you could have done otherwise might not mean the same as the ability to make a choice that is undetermined. Whether they believe this would also decide whether they think that free will implies the ability to make an undetermined choice, and thus if they agree with Conway and Kochen about free will.

So, for compatibilists, it is not obvious to say that the ability to make an undetermined choice follows from free will, but they are an exception in this. However, as compatibilism is a popular position in the modern debate on free will, this is something to be mindful of when using a certain definition of free will. Additionally, not all agree that an undetermined choice leaves the possibility for free will (see luck objection in section 2.3). However, it seems impossible to satisfy everyone in the free will debate with one definition of free will, because the meaning of free will means depends on your position in the debate. So, you have to choose between all the different definitions at some point. The CHDO condition, in the sense that you have options to choose from, is an important condition for free will in the overall debate. This shows that they have not defined free will in a very controversial way here. So, many would agree that it is reasonable for Conway and Kochen to say that the FWT relates to free will because the experimenter's ability to make a choice that is undetermined by past events follows from their free will.

6 Free will for a human and a particle

In the former chapter, it was discussed how Conway and Kochen connect the FWT to free will when they say the free will assumption follows from the free will of the experimenter. But is it enough that an important assumption of a theorem follows from free will, to call it the 'Free Will Theorem'? It seems that the main result of the theorem should also be related to free will for this to be a good title.

Conway and Kochen state the main result of the FWT as the surprising consequence that a particle is free in the same sense as the experimenter [6] (p. 228).

Our theorem is the surprising consequence that particle a's response must be free in exactly the same sense, that it is not a function of what has happened earlier (with respect to any inertial frame).

So, both the experimenter and the particle have freedom in the sense that they have the ability to make a choice that is undetermined by past events. Does this type of freedom from the FWT also result in free will for the particle? To see whether the particle in the FWT has free will, I will first examine what Conway and Kochen say about this. Then we will see how this compares to the free will of the experimenter, which was discussed in Chapter 5. It then follows from this comparison what Conway and Kochen could mean by ‘free will’.

6.1 The free will of a particle

In the paper about the Strong FWT [6] (p. 230) Conway and Kochen state that they also speak of free will in case of a particle.

Some readers may object to our use of the term “free will” to describe the indeterminism of particle responses. Our provocative ascription of free will to elementary particles is deliberate, since our theorem asserts that if experimenters have a certain freedom, then particles have exactly the same kind of freedom.

Conway and Kochen admit that their use of the term ‘free will’ to describe the undetermined response of a particle is provocative. However, I will not look at how the use of the term ‘free will’ to describe a particle is in itself provocative, but at how the use of this term can be misleading within the FWT. For this, I will first distinguish how ‘free will’ is used for a human and particle in the FWT.

Figure 3 shows the relation of the FWT to free will for both a human and a particle. The bottom black arrow represents the result that the FWT aims to prove. The assumption that for the experimenter the ability to make a choice that is not determined by past events follows from free will (as discussed in section 5.3) is shown in the arrow on the left. It can clearly be seen in Figure 3 that the FWT does not directly prove free will for particles. Only in the case that arrow 2 is a correct implication, would it be possible for Conway and Kochen to argue that the FWT proves the ‘free will’ of a particle.

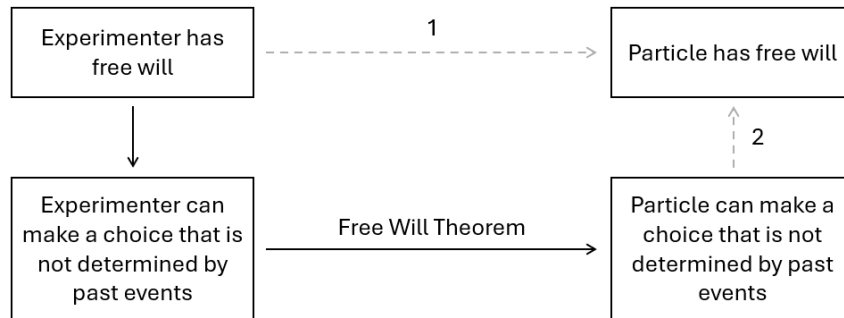


Figure 3: A visual representation of the implication of the FWT in relation to free will.

Since, in their paper, Conway and Kochen speak of the free will of a particle, it seems that they would say the implication in arrow 2 is true. If this implication is true it means that if a particle can make a choice that is undetermined by past events this particle has free will. Then in the case of

the particle, the ability to make an undetermined choice is a sufficient condition for free will. This might sound reasonable, but it gets when confusing considering that free will for the experimenter was not defined in the same way in the FWT. For a human, the ability to make a choice that is undetermined by past events was regarded as only being a part of their free will. For humans, this ability is a necessary condition for free will¹⁷, but not a sufficient one. This shows ambiguity in the usage of the term ‘free will’ for humans and particles in the FWT. It seems that they mean something different when speaking of the free will of a particle than what they mean by the free will of a human.

6.2 Two interpretations of free will

To see how the FWT relates to free will, it is desirable to find a way to avoid ambiguity in the use of ‘free will’ for particles and humans. For a consistent definition of ‘free will’ for both the human and the particle throughout the FWT there are two options.

1. Free will is defined by the condition that you have the ability to make a choice that is undetermined by what has happened in the past.
2. Free will consists of multiple conditions of which one is the condition that you have the ability to make a choice that is undetermined by the past.

Only option 1 allows for implication 2 in Figure 3 to be true. So, only in option 1, the FWT proves free will for particles. This is the result that Conway and Kochen seem to aim for when they state that their theorem relates to free will for a particle.¹⁸ Their emphasis on free will, especially for particles, indicates that they want the FWT to relate strongly to free will. So, option 1 gives an interesting result of the FWT for Conway and Kochen, but this comes at the cost of using a very narrow definition of free will. To define free will only by the ability to make an undetermined choice or equally the CHDO condition, is at the least very unusual in philosophy. Though disagreeing on the exact meaning of free will, philosophers from most positions in the free will debate agree that agency plays at least some role in free will. So, free will as the ability to make an undetermined choice is not a version of free will that is very meaningful in the philosophical debate. So, in the case of option 1 the human and particle would both have the same free will, but for philosophy, this would not be a very relevant result.

Option 2 allows for a broader notion of free will for the experimenter, where the ability to make an undetermined choice is only part of free will. This aligns much better with free will in philosophy. So, that the definition of free will is much more narrow than in philosophy disappears, in contrast to option 1. But this comes at a cost, as in option 2, the implication in arrow 2 in Figure 3 is not true. So, now the free will assumption in the FWT can follow from the experimenter’s free will in a meaningful way, but then the FWT does not prove free will for the particle. The particle only gains the ability to make an undetermined choice, which is supposed to be a part of the free will of the experimenter. It can then even be questioned whether this ability, that the particle gains from the FWT, is related to free will at all. So, in option 2, the FWT does not prove free will for a particle and Conway and Kochen cannot speak of the particle’s free will as a result of the FWT.

So, the relation of the FWT to free will turns out to be more difficult when regarding also the particle. Either the FWT does prove free will for the particle and human but uses a version of free

¹⁷Even this is arguable, see section 5.3

¹⁸Unless they mean something different by the free will of a particle and the free will of a human. But they say that the FWT proves free will for particles, without giving any clarification on this being a different use of the term ‘free will’ than for the experimenter.

will so stripped down, to only the ability to make an undetermined choice, that it is now a separate kind of free will from that in philosophy. Otherwise, the FWT does allow for a broader and stronger notion of free will, but then the FWT does not prove this type of free will for the particle. This shows that either way, the FWT does not prove free will for particles in a philosophically meaningful way. This might make the name ‘Free Will Theorem’ less relevant. However, the surprising result of the FWT, that it is supposed to prove that if a human can make an undetermined choice so can a particle, remains whether or not this can be called free will.

7 Free will, the undetermined choice and indeterminism

As discussed in the former chapter, it turns out that the relation of the FWT to free will is troublesome. The version of freedom that Conway and Kochen use in the theorem, the ability to make an undetermined choice, is not necessarily the same as having free will. Yet, they choose only this specific type of freedom for what the FWT aims to prove. In this chapter, I will look at possible reasons for choosing this type of freedom and the consequences of making this choice. For this, I will examine Conway and Kochen’s statements about free will in the FWT.

7.1 The experimenter’s freedom

Conway and Kochen first relate the FWT to free will through the free will assumption, which states that an experimenter can make an undetermined choice (as discussed in section 5.2). They mention that they treat the freedom to choose experimental settings as a tacit assumption in the MIN axiom [6]. This might sound like a very reasonable assumption when thinking about a physics experiment. With an experiment, a physicist tries to strengthen a theory by testing if it works in different environments. To test many possible environments, measurements are performed in different experimental settings. If all the settings that the experimenter can choose are predetermined, there might be other settings that the experimenter can never choose. In that case, how can you ever know if a theory works for all settings? So, the belief that an experimenter is free to choose settings, is very useful for making empirical investigation possible. This might explain why Conway and Kochen treat the freedom of the experimenter as an implicit assumption, but it is not an argument for believing the free will assumption.

But this is not a reason for the specific type of freedom in the FWT either, because being able to choose settings for an experiment is not necessarily the same as having the ability to make an undetermined choice.¹⁹ It seems that for experimental methods to be fruitful, the experimenter should be able to choose freely in at least some sense, that is in the sense that the choice of the experimenter should not be determined by whatever determines the outcome of the experiment. This still allows for the choice of the setting of the experiment to be determined by separate information from the past. The freedom that Conway and Kochen assume is the ability to make a choice undetermined by everything that has happened in the past. This is a broader assumption. So, the necessity of some freedom to make experimental methods possible is not an explanation for the specific type of freedom that Conway and Kochen assume for the experimenter in the free will assumption, as these are different types of freedom.

Conway and Kochen also want to be able to replace the assumption that the experimenter can make an undetermined choice about their settings with a broader assumption of free will. Free will in a broad sense also encompasses much more than just the freedom that is necessary for experimental

¹⁹t Hooft [36] argues that the freedom necessary for physics models is not the same as the freedom that Conway and Kochen assume.

methods. So, the necessity of being able to choose measurement settings freely for the usefulness of the experimental method is not an argument to believe that any of these broader versions of freedom should be assumed.

It was also shown that why Conway and Kochen use their specific type of freedom, as the ability to make an undetermined choice, is not by the freedom an experimenter desires. It was also discussed in section 5.2 that free will in philosophy is more than the small part of freedom that Conway and Kochen choose to prove the FWT. So, it seems that the reason for this specific type of freedom does not stem from the experimenter. This leaves the question of why they would choose only such a specific part of the free will of the experimenter open.

7.2 Variations of free will for a particle

To define freedom as the ability to make an undetermined choice, without necessarily agency, seems like a useful definition for talking about the free will of particles. Since ascribing agency to particles seems much harder than arguing that particles can act in an undermined way. Conway and Kochen use a type of freedom that is much more applicable to a particle than some versions of free will from philosophy might be. Might the type of freedom they use then originate from the type of freedom that is useful for describing the behaviour of particles?

In the original paper on the FWT [5] (p. 1466) Conway and Kochen state that they define ‘free will’ as the opposite of determinism.

We have defined “free will” to be the opposite of “determinism” despite the fact that since Hume some philosophers have tried to reconcile the two notions – a position called compatibilism. In our view this position arose only because all the physics known in Hume’s day was deterministic, and it has now been outmoded for almost a century by the development of quantum mechanics. However, for the purposes of our paper, we can bypass this hoary discussion, simply by saying that the only kind of free will we are discussing, for both experimenters and particles, is the active kind of free will that can actually affect the future, rather than the compatibilists’ passive variety that does not.

By defining free will as indeterminism, they define free will as in option 1 from Chapter 6.2, which is free will that is defined solely by the ability to make an undetermined choice. A similar view comes forward when they mention that they describe the ‘indeterminism of particle responses’ as free will. This view of indeterminism as freedom works when you try to explain indeterministic particle responses as free actions. However, this is not a logical definition of free will when looking at humans or the philosophical free will debate. This can be a reason to believe, that they started with some freedom a particle could have, and then extended this towards the human.

A reason for Conway and Kochen to also describe the free will of humans in terms of a type of freedom that particles have might be their reductionist beliefs about free will that they describe in the paper. They say that it is natural to assume that the freedom of particles is an explanation of the freedom of humans [6] (p. 230).

Our provocative ascription of free will to elementary particles is deliberate, since our theorem asserts that if experimenters have a certain freedom, then particles have exactly the same kind of freedom. Indeed, it is natural to suppose that this latter freedom is the ultimate explanation of our own.

But the fact that it sounds natural to them, should not be a reason to describe the freedom of the experimenter in terms of the freedom of the particle. Especially for a theorem that proves if

humans have certain freedom, then so do particles, it is not desirable to choose a version of freedom for the human, based on what would work for the particle. This seems like reverse engineering the theorem. Unless they would have a good reason to describe the freedom of humans in exactly the terms that are so useful to describe the freedom of particles. But unfortunately making the comparison to philosophy, this is not a logical definition of freedom for a human. In the next part, I will further discuss how Conway and Kochen's version of freedom can become problematic with regard to randomness, especially in the case of the experimenter.

7.3 Undetermined choice and randomness

When Conway and Kochen talk about free will as indeterminism, it raises the question of what makes an indeterministic particle response free and not random. Conway and Kochen justify why they do not describe the behaviour of their particles as random [5] (p. 1466).

Although we find ourselves unable to give an operational definition of either “free” or “random,” we have managed to distinguish between them in one very special context. Bassi and Ghirardi remark that it follows from Gisin's theorem that their “hits” must involve a stochastic element in order to make the GRW theory relativistically invariant (while remaining causal). We showed in Section 10 that this is not enough, even if it is combined with non-local correlations (that can be arbitrarily strong provided they do not grossly violate causality by informing b about x, y, z or a about w). What the hits really need is some freedom (to be precise, they must be at least semi-free). It is for reasons including these that we prefer to describe our particles' behaviour as “free” rather than “random,” “stochastic,” or “indeterminate.”

Whether this is an acceptable reason to call the behaviour of particles random, is a question that will not be answered here. But even if this is a valid reason for the behaviour of the particles to not be random, this argument cannot be applied in the same way to the experimenters. The question of whether the choice of the experimenter is random remains. Conway and Kochen state that the choice of the experimenters is also free, and not random, as they make clear in the Strong FWT [6] (p. 230).

Although the FWT suggests to us that determinism is not a viable option, it nevertheless enables us to agree with Einstein that “God does not play dice with the Universe.” In the present state of knowledge, it is certainly beyond our capabilities to understand the connection between the free decisions of particles and humans, but the free will of neither of these is accounted for by mere randomness.

They give no explicit argument for why the choice of the experimenter is free and not random. Without a convincing argument, one can wonder why an undetermined choice of an experimenter would be free instead of random.

A choice that is in no way related to what has happened in the past seems to be random. Yet, Conway and Kochen would allow this type of choice for their experimenters. Only a very small part, if any, of the free choices a human makes seem to be made without any reason. Most of our choices are, at least in some way, based on considerations about what has happened in the past. But this is not a requirement for Conway and Kochen's free choice. In the view of Spinozic realism it is not even possible for a choice that is undetermined by what has happened in the past to have reasons at all. Because in this view, any change in the physical world has to be determined within the physical world. So, if there were reasons that determine a choice that causes physical action, these reasons

are also physical. Then a choice that has reasons cannot be undetermined. Therefore, in the view of Spinozic realism we cannot have reasons for any undetermined choice.²⁰ Again, a choice that has no reasons at all, seems more random than free. So, it is not obvious that an undetermined choice is the result of human freedom and not randomness.

But even if it is granted that the choice of the experimenter can be in some way influenced by what has happened in the past, but is not determined by this, the question remains whether it is really a free choice. This case would allow the choice to be made under some part of agential control. But in philosophy, it is still an open debate how free will can exist in an indeterministic world, or based on undetermined choices. It can even be argued that any undetermined choice is random, and therefore always excludes the possibility of agency. This can be done by the ‘luck objection’ (as discussed in section 2.2), which is the objection that whether a choice that is caused indeterministically is just a matter of luck. Thus, the choice would be random and not free. Whether this argument works is still an open discussion, of which the answer would impact whether Conway and Kochen can call an undetermined choice a free choice.

So, being undetermined by what has happened in the past as the only condition of freedom, leads to the objection that this freedom might be random. This could be a reason to object to interpretation 1 of free will in section 6.2, which states that free will is solely defined by the ability to make an undetermined choice. In the case of option 2, in which only being undetermined is a part of free will, agency is allowed as a part of the free will of the experimenter. But in philosophy, it is not even sure whether this type of free will is possible, as it can be argued that an undetermined choice is random and therefore excludes agency. However, this is still an open debate, that relates to the question of what versions of free will would be possible at all in an indeterministic world. So, it relates to the question of whether libertarianism is possible at all. Conway and Kochen use a type of freedom that aligns well with the libertarian account of free will which includes indeterminism, but can also be criticized by opponents of this theory.²¹

So, overall, the version of freedom that Conway and Kochen choose can cause some difficulties when applying this type of freedom to the experimenter. Because if the choice is completely undetermined by anything in the past, it seems like a random event. If the choice can be in some way influenced by reasons, which is not a necessary condition for Conway and Kochen, but also not explicitly excluded by them, then the problem lifts as the choice does not have to be random (although some philosophers might still argue this).

7.4 Free will in an indeterministic world

Conway and Kochen describe their free will as indeterminism, for particle this description seems to work well. However, in the philosophical debate on free will, this is a big claim. In this debate, it is still an open question how free will relates to determinism or indeterminism. To define free will as the opposite of determinism is to say that free will and determinism are incompatible. So, to define free will as indeterminism can be seen as an incompatibilist assumption. However, Conway and Kochen’s definition of free will as indeterminism is only once mentioned explicitly, so I will first examine what exactly they mean by this.

Conway and Kochen give no exact definition of what they mean by indeterminism, but they describe that a non-deterministic world makes it impossible to return to a clockwork universe [6] (p. 230).²²

²⁰This leaves the possibility of feeling like we have ‘reasons’ for a choice, while these are really an illusion. This does not change the situation as the choice would still be physically determined.

²¹An overview of possibilities for incompatibilist free will can be found in [14].

²²This quote is now used to get an idea of how Conway and Kochen view (in)determinism, regardless of its credibility.

The import of the FWT is that it is not only current quantum theory, but the world itself that is non-deterministic, so that no future theory can return us to a clockwork universe.

As a clockwork universe, the world would be a perfect machine that acts according to the laws of nature, making it fully predictable. It seems that this is what they view as a deterministic world. But, they also use (in)determinism as a property to describe theories, like they do here for quantum mechanics.

Conway and Kochen's indeterministic particle response resembles the version of determinism by Earman that was established as relevant to the free will debate in section 2.1. Although determinism or indeterminism, as defined by Earman, is a property of the world, this resemblance remains. An undetermined choice is a choice that is not a function of, so not uniquely determined by, what has happened in the past. Similarly in a world that is indeterministic by Earman, not all states are uniquely determined, but even not by a single state. Still, indeterminism by Earman can not be used to describe a single choice.

Now if we ignore that Conway and Kochen claimed free will to be indeterminism, we can look back at their more specified definition of freedom that is used in the FWT. They define the freedom of the experimenter as the ability to make a choice undetermined by what has happened in the past. This definition might have an impact on the indeterminism of the world. Still, in the view of Spinozic naturalism, the experimenter is a part of the deterministic natural world and their action is determined by natural law. As defined by Earman, a deterministic world means a world for which all worlds with the same natural laws, if they are in the same state at some time, they are in the same state at all times. An indeterministic world is the opposite of a deterministic world, meaning that even if two possible worlds are in the same state at some time, they can still be in a different state at another time. So, in an indeterministic world, a choice not determined by past events can exist. In a deterministic world, an undetermined choice is not possible. So, a world has to be indeterministic, if an undetermined choice exists in it. This shows that by assumption (3a), which states that an experimenter can make an undetermined choice, Conway and Kochen also assume an indeterministic world as a starting point for their theorem.

Yet, not everyone agrees that it is only possible to make an undetermined choice in an indeterministic world. These are the compatibilists, who believe you could have done otherwise, even in a deterministic world, e.g. Lewis. So, they believe your choice does not have to be determined even if the world is deterministic. Interpreting the undetermined choice of the FWT in such a way would allow the free will assumption to exist in both a deterministic and indeterministic world. Then it would not follow from the free will assumption that the world is indeterministic.

8 The Free Will Theorem as an argument in the free will debate

The main problem in the debate on free will nowadays is whether, and how, free will and determinism are compatible. I will now examine whether the FWT can be used as an argument for any of the major positions in this debate. I will first discuss Conway and Kochen's own views about the free will debate and see how these can be realized in section 8.1. Then to see more clearly how the FWT relates to the free will debate, I will discuss a reformulation of the FWT by Cator and Landsman [37] in section 8.2. Possible solutions to the problems that arise when trying to use the FWT as an argument in the debate on free will follow in section 8.3.

8.1 The Free Will Theorem as an argument for libertarianism

To see what Conway and Kochen's position in the debate on free will is, I will first look at the claims Conway and Kochen make about free will, determinism and compatibilism in their papers on the FWT as a guideline.

Conway and Kochen assume an indeterministic world for the FWT with the free will assumption. They also claim that the FWT expands the indeterminism of quantum mechanics, to the entire world. So, any position they argue for will not be determinism.

Conway and Kochen also argue, when they define free will as indeterminism, that the position of compatibilism can be bypassed by the theorem. They say they only regard the active version of free will that can change the world, not the passive kind compatibilists use. They do not specify which exact versions of compatibilism they mean by this. But they do show they are not convinced by compatibilism, as they think that 'this position arose only because all the physics known in Hume's day was deterministic, and it has now been outmoded for almost a century by the development of quantum mechanics' [5]. Their dislike towards determinism and compatibilism, while assuming free will, leaves only one position for them in the free will debate; libertarianism.

However, the FWT itself does not give a strong argument for libertarianism. For this, they would have to prove that free will exists and the world is indeterministic. Since free will is assumed in the FWT, the FWT cannot be used as an argument for free will. Only those who believe in free will agree with this assumption, these are libertarianist and compatibilists. The free will theorem might then still be useful for proving that if free will exists the world is indeterministic. This way it can still be useful in favour of libertarianism, by weakening their compatibilist opponents, as it shows that free will does not allow for determinism. But this would only work for the compatibilists who agree with the type of freedom that is used in the FWT.

It is not all compatibilists that agree with the free will assumption that is made in the FWT, for example, new compatibilists would argue that whether a choice is determined is not important for free will. Conway and Kochen exclude all compatibilism at once from their theory, by saying that they do not regard the compatibilist's 'passive' type of free will. But what they mean by a 'passive' type of free will and at which exact compatibilist theories they aim is not specified. So, there is still some room for interpretation. It might be, despite Conway and Kochen's statement, there are compatibilists who have the same understanding of free will as they do. The theorem can then be used to argue against these compatibilists, who view the ability to make an undetermined choice as a part of free will. For this, the FWT would have to show that the free will assumption leads to indeterminism and therefore an incompatibility with determinism.

The result of the FWT is supposed to be that particles can respond in a way that is not determined by what has happened in the past. This then indeed results in an indeterministic world. So, you could say that indeed the FWT proves that the free will assumption excludes determinism. The problem with this reasoning is that the free will assumption, that a human can choose in a way undetermined by what has happened in the past, also already implies an indeterministic world (as discussed in section 7.4). So, they aim to prove indeterminism by assuming indeterminism in the first place. This circularity is shown in Figure 4. This has been pointed out before, e.g. Wüthrich [24] argues that 'as far as [the Free Will Theorem] is concerned, it's indeterminism in, indeterminism out.'

So, to use the FWT as an argument against certain types of compatibilism, with to indirectly strengthen the position of libertarianism, a way needs to be found to avoid this circularity. But first, despite Conway and Kochen's views, I will discuss whether there are other possibilities to use the FWT in the discussion on free will and determinism.

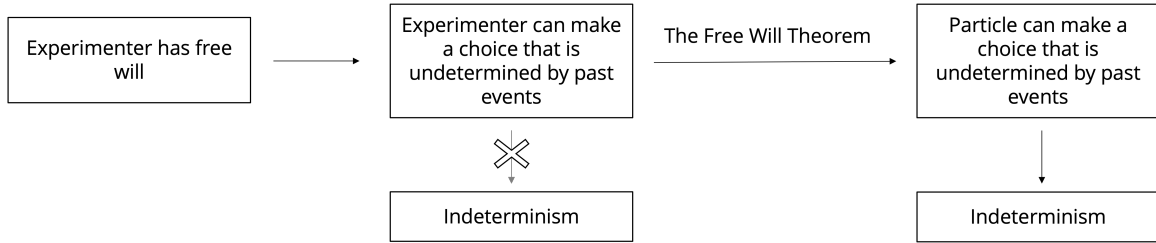


Figure 4: A visual representation of the circularity that should be avoided to be able to use the FWT as an argument against certain types of compatibilism.

8.2 Reformulation of the Free Will Theorem

To get an overview of what the FWT proves, a reformulation of the FWT will be discussed that was proposed by Landsman and Cator [37]. In this reformulation, the free will of the experimenter is no longer among the initial assumptions. But instead, the FWT shows a contradiction between the free will of the experimenter and the determinism of the particle. In this reformulation, the axioms of the theorem are replaced by the following axioms. These are described very loosely here just for the purpose of getting an idea of what the reformulation does, but the actual descriptions are very mathematical and more precise as described in [37].

- **FREEDOM:** The experimenter's choice is independent of outside variables and the choice of the other experimenter. (Similar to Part a of MIN axiom from section 6.2)
- **DETERMINISM:** For both particles, there exists a function, that determines the measurement outcome (some form of 1, 0, 1) as a function of the corresponding measurement settings and outer variables, for all possible options.
- **NATURE:** The result of a squared spin measurement in three orthogonal directions of a spin 1 particle is 1, 0, 1 and the particles in the experiment have a perfect correlation. (Similar to SPIN and TWIN axiom)
- **LOCALITY:** The result of the measurement of particle a is independent of the measurement directions of particle b and the other way around. (Similar to Part b of MIN axiom from section 6.2)

The result of the FWT is then that NATURE and LOCALITY contradicts with FREEDOM and DETERMINISM. So, if the physics axioms are accepted, then FREEDOM and DETERMINISM cannot both hold. Then if FREEDOM is accepted, DETERMINISM cannot hold and this leads to the same result as the FWT. The other way around, if DETERMINISM is assumed, FREEDOM cannot hold. This way, the FWT can also be used as an argument that shows determinism excludes free will. This can be done if you assume a deterministic world, the choice of a particle is a function of the past. Then in a deterministic world, an experimenter also cannot choose their settings undetermined by what has happened in the past by the FWT. The FWT then proves that in a deterministic world, the experimenter's choice is determined by what has happened in the past.²³

²³The FWT then gets a similar meaning to the Consequence Argument, but instead of showing you could not have done otherwise in a deterministic world, the FWT shows that a human cannot make an undetermined choice in a deterministic world.

But this argument only works if indeed every response by a particle in a deterministic world is also determined, otherwise, the connection to determinism would be lost.

Now coming back to finding a way to use the free will theorem as an argument against compatibilism, this reformulation shows that in this case the FWT would be an indirect support for both libertarianism and hard determinism. However, the circularity discussed before still needs to be solved, as it remains also in this reformulation. FREEDOM states that the experimenter can choose independently of all outside variables, so the choice can not be determined by these outer conditions. So, just like in the original FWT, the experimenter can make a choice that is not determined by what has happened in the past. Then if choosing to accept FREEDOM, you again assume an indeterministic world.

When assuming DETERMINISM, the FWT can also be used as an argument against compatibilism. The FWT then proves that determinism excludes free will and thus compatibilism with this version of free will cannot exist in a deterministic world. But the same circularity as before is also a problem in this case. The circularity now arises when determinism, without the FWT, implies that an experimenter cannot make a choice that is undetermined by the past, thus that free will cannot exist. Thus, again the idea that an undetermined choice for the experimenter can only exist in an indeterministic world causes a circularity. So, if this circularity can be avoided, the FWT can be used as an argument against certain types of compatibilism, and it is then an indirect argument equally for libertarianism or hard determinism.

8.3 Solutions to circularity

To use the FWT as an argument against certain types of compatibilism, the circularity in Figure 4 needs to be avoided. This circularity disappears for those who believe that an experimenter can make a choice undetermined by what has happened in the past, both in a deterministic and indeterministic world. But for the result of an indeterministic world to maintain, they should also believe that the ability of a particle to make a choice that is undetermined by what has happened in the past can only exist in an indeterministic world. This way the FWT can be used as a necessary step to prove that if a human has free will, the world is indeterministic. Also, it is still necessary for this reasoning that the compatibilist agrees with the version of free will in the FWT, in which it follows from the free will of a human, that they can make an undetermined choice. So, what we are looking for is compatibilist theories in which it is a necessary condition of free will that you can make an undetermined choice, but where this ability to make an undetermined choice does not imply an indeterministic world.

Landsman showed that the type of freedom in the FWT is the same as how the local miracle compatibilist defines that you could have done otherwise [25]. In the version of free will that a local miracle compatibilist uses, a person could have done otherwise in a deterministic world. So for them, the undetermined choice of the experimenter does not imply indeterminism. In this case, the FWT proves, without circularity, that a local miracle compatibilist's version of free will results in indeterminism and is not compatible with determinism. So, the FWT shows that a local miracle compatibilist notion of free will does not allow for compatibilism. But it needs to be mentioned that in this case, the ability of a particle to make an undetermined choice, should in the local miracle compatibilist view not be possible in a deterministic world. So, in this case, the ability to make an undetermined choice has a different meaning for the experimenter and the particle.

For the interpretation of freedom that is necessary to use the FWT as an argument against compatibilism, the freedom of the particle and the human should have a different impact on the determinism of the world. A human should be able to have this type of free will in a deterministic world, while if a particle has the same type of freedom it would imply indeterminism. The human

and particle then both have different abilities to act freely. This would defeat the original purpose of the FWT, which was to show that a human and a particle have exactly the same type of freedom.²⁴ So, it shows that for the FWT to have a meaningful philosophical result, the freedom of the human and the particle have to be not exactly the same.

Now if we let go of the assumption of Spinozic naturalism, that said human behaviour falls under natural law as well, this might allow for additional versions of compatibilism to be discarded. A theory of mind-body dualism, where the mind can act freely outside of the law-governed deterministic natural world, could also result in a version of compatibilism that satisfies what we are looking for. But this would require a sufficient definition of determinism for the natural world, that would allow for an immaterial mind. The FWT could then show that if there is an immaterial mind and free will, the world needs to be indeterministic. If then the freedom of the experimenter and particle is in some way the same, as the FWT aims to show, the FWT might hint that human freedom is possible because of the indeterminism of particles in this case.

9 Conclusion

The FWT aims to prove that if an experimenter has a certain type of freedom, then a particle has this as well. This specific type of freedom is the ability to make a choice that is undermined by what has happened in the past. In philosophy, free will usually entails more than just the ability to make an undetermined choice or it is even completely different. Yet, Conway and Kochen refer to their theorem as the ‘Free Will Theorem’.

Conway and Kochen speak of free will for both the particle and the human in their papers on the FWT. This is only possible if they say free will is defined solely as the ability to make an undetermined choice. Then both the particle and human have free will by the FWT. But a very narrow notion of free will is used that is not at all common in the philosophical debate. However, Conway and Kochen can also use a broader version of free will and accept that the FWT does not prove free will for the particle. This free will is broader than but does include, the ability to make a choice undetermined by past events. Then the theorem uses a more meaningful version of free will for the experimenter from which the free will assumption follows but it does not prove this version of free will for a particle. In either of these two options for the definition of free will, the FWT does not prove free will for the particle in a way that is meaningful to the philosophical free will debate.

Regardless of what definition of free will Conway and Kochen use exactly, the theorem aims to prove something about the ability to make an undetermined choice. Why did they choose this version of freedom for a theorem about free will? Freedom as the ability to make an undetermined choice, or as Conway and Kochen describe it free will as indeterminism, is a version of freedom very hospitable for particles to have. However, it turns out that this definition is not so obvious to use for the experimenter. In the first place, it is not common in philosophy and also does not follow from the freedom that is necessary for experimental methods. Additionally, it can even lead to objections regarding randomness for the choice of the experimenter. So, they pick out a part of human free will for the theorem, especially hospitable to particles. A reason for this might be their reductionist assumption that this type of freedom for humans follows directly from the freedom of particles.

Despite its difficult relation to free will, the FWT can be used as an argument in the free will debate. The FWT be an argument against only specific types of compatibilism, and can thereby strengthen libertarianism or hard determinism. This can be done when using the result of the

²⁴Or the freedom would be the same but it would have a different impact on the physical world when enacted by a human and a particle. Which would leave a difference in the results that the particle and the experimenter can achieve with this freedom. So, their freedom would still be different when it comes to acting in the physical world.

FWT, which is that particles are undetermined in their responses, as proof of indeterminism. Yet, this leads to a circularity, because the free will assumption in the FWT already in itself implies indeterminism. This circularity can be avoided for versions of free will for the experimenter that do not imply indeterminism. So, the FWT can be used as an argument against compatibilists who share this view of free will. These are specifically, compatibilists who agree that for a human it is a necessary condition for free will to have the ability to make a choice that is undetermined by what has happened in the past, but this ability does not imply an indeterministic world. Yet, for a particle, the ability to make an undetermined choice should imply indeterminism. So, in any case, for this to work, the undetermined choice of the human should not imply indeterminism, while the undetermined choice of the particle should.

Landsman proposed a version of compatibilism that can be disproved by the FWT. He shows that the ability to make an undetermined choice in the FWT is the same as how you could have done otherwise in the local miracle compatibilist's version of free will. Thus, the FWT can be used as an argument against local miracle compatibilism, on the condition that the free response of the particle in this interpretation still implies an indeterministic world.

Still, this leaves new compatibilists (and possibly more compatibilists) unharmed by the FWT, so many versions of compatibilism remain possible. However, there are more contemporary compatibilist theories, which could possibly be ruled out by the FWT in the same way as local miracle compatibilism. Figuring out which specific versions of compatibilism the FWT can apply to, could be a topic of further research. The FWT can play a role in the debate of free will and determinism by disproving these versions of compatibilism. Additionally, there might be possibilities for using the FWT theorem as an argument against mind-body dualism. Other possible ways to use the FWT in the debate on free will, that were overlooked here, might also be found.

So, it was shown that to relate the FWT to free will meaningfully, it has to be accepted that the FWT does not prove free will for the particle, but only a certain type of freedom. Then for the FWT to be an argument in the free will debate, the free will of the particle and human should have a different effect on determinism. So, to make the FWT relevant in the philosophical debate, the freedom of the particle is first of all not actual free will, and secondly not of exactly the same type as the freedom of the experimenter. Thus, for using the FWT theorem as an argument in philosophical debate, the result of the FWT as Conway and Kochen put it 'Our theorem is the surprising consequence that particle a's response must be free in exactly the same sense [as experimenter A's choice]'[6], cannot hold. But then in showing that certain compatibilist theories, where the freedom of the human and particle are not the same, need to be rejected, the FWT could hint at theories in which the freedom of the human and the particle are not that different after all.

In the FWT many difficulties arise in investigating the conceptual philosophical problem of free will through the science of physics and mathematics. It shows this is difficult in the sense of finding the right philosophical definitions and translating these to physics for finding a theorem. These difficulties that have been shown to occur in the FWT can be something to keep in mind for other physicists, who in the future would also want to make a difference in the debate on free will and determinism. Hopefully, there will still be those who dare to try this. Since, despite saying that some questions about free will require the conceptual analysis of philosophy instead of empirical investigation, Slors, de Bruin and Strijbos mention the following in the *philosophy of mind, brain and behaviour* [12].

When it comes to determinism versus the possibility to do otherwise, this is mainly a philosophical issue. Scientists who would be in a position to contribute to this issue are those that deal with the possibly deterministic nature of our world, i.e. theoretical physicists.

References

- [1] B. W. Libet, “Do we have free will?”, *Journal of Consciousness Studies* **6**, 47–57 (1999).
- [2] G. Gomes, “The timing of conscious experience: a critical review and reinterpretation of libet’s research”, *Consciousness and Cognition* **7**, 559–595 (1998).
- [3] D. M. Wegner, *The illusion of conscious will* (MIT press, 2017).
- [4] P. V. Inwagen, “The incompatibility of free will and determinism”, *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition* **27**, 185–199 (1975).
- [5] J. H. Conway and S. Kochen, “The Free Will Theorem”, *Foundations of Physics* **36**, 1441–1473 (2006).
- [6] J. H. Conway and S. Kochen, “The strong free will theorem”, *Notices of the American Mathematical Society* **56**, 226–232 (2009).
- [7] A. Hüttemann, *Free will in a law-governed universe*, September 2024, Manuscript (2024).
- [8] J. Butterfield, “Determinism and indeterminism”, *Routledge encyclopedia of philosophy* **3**, 3339 (2005).
- [9] J. Earman, *A primer on determinism* (Reidel Publishing Company, 1986).
- [10] M. Van Strien, “Was physics ever deterministic? the historical basis of determinism and the image of classical physics”, *The European Physical Journal H* **46**, 8 (2021).
- [11] A. Hüttemann, “Determinismus - eine empirische these”, *Zeitschrift für Philosophische Forschung* **76**, 479–509 (2022).
- [12] M. Slors, L. de Bruin, and D. Strijbos, *Philosophy of mind, brain and behaviour* (Boom, 2015).
- [13] T. O’Connor and C. Franklin, “Free Will”, in *The Stanford encyclopedia of philosophy*, edited by E. N. Zalta and U. Nodelman, Winter 2022 (Metaphysics Research Lab, Stanford University, 2022).
- [14] R. Clarke, J. Capes, and P. Swenson, “Incompatibilist (Nondeterministic) Theories of Free Will”, in *The Stanford encyclopedia of philosophy*, edited by E. N. Zalta, Fall 2021 (Metaphysics Research Lab, Stanford University, 2021).
- [15] D. Widerker, “On the luck objection to libertarianism”, in *Agency, freedom, and moral responsibility*, edited by A. Buckareff, C. Moya, and S. Rosell (Palgrave-Macmillan, 2015), pp. 94–115.
- [16] D. Moore, “Libertarian free will and the physical indeterminism luck objection”, *Philosophia* **50**, 1–24 (2021).
- [17] M. McKenna and D. J. Coates, “Compatibilism”, in *The Stanford encyclopedia of philosophy*, edited by E. N. Zalta and U. Nodelman, Summer 2024 (Metaphysics Research Lab, Stanford University, 2024).
- [18] P. F. Strawson, “Freedom and resentment”, *Proceedings of the British Academy* **48**, 187–211 (1962).
- [19] H. G. Frankfurt, “Alternate possibilities and moral responsibility”, *Journal of Philosophy* **66**, 829–839 (1969).
- [20] D. Lewis, “Are we free to break the laws?”, *Theoria* **47**, 113–21 (1981).
- [21] H. Beebe, “Local miracle compatibilism”, *Noûs* **37**, 258–277 (2003).

- [22] S. Kochen and E. P. Specker, “The problem of hidden variables in quantum mechanics”, *Journal of Mathematics and Mechanics* **17**, 59–87 (1967).
- [23] A. Einstein, B. Podolsky, and N. Rosen, “Can quantum-mechanical description of physical reality be considered complete?”, *Phys. Rev.* **47**, 777–780 (1935).
- [24] C. Wuthrich, “Can the world be shown to be indeterministic after all?”, in *Probabilities in physics*, edited by C. Beisbart and S. Hartmann (Oxford University Press, 2011), pp. 365–389.
- [25] K. Landsman, “On the notion of free will in the free will theorem”, *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* **57**, 98–103 (2017).
- [26] R. Tumulka, “Comment on “the free will theorem””, *Foundations of Physics* **37**, 186–197 (2007).
- [27] A. Bassi and G. Ghirardi, “The conway-kochen argument and relativistic grw models”, *Foundations of Physics* **37**, 169–185 (2007).
- [28] S. Goldstein, D. Tausk, R. Tumulka, and N. Zanghi, “What does the free will theorem actually prove?”, *Notices of the American Mathematical Society*, 1451–1453 (2010).
- [29] J. S. Bell, “On the problem of hidden variables in quantum mechanics”, *Reviews of Modern Physics* **38**, 447–452 (1966).
- [30] A. Elby, “Nonlocality and gleason’s lemma. part 2. stochastic theories”, *Foundations of Physics* **20**, 1389–1397 (1990).
- [31] D. Hemmick, “Hidden variables and nonlocality in quantum mechanics”, (2004).
- [32] H. R. Brown and G. Svetlichny, “Nonlocality and gleason’s lemma. part i. deterministic theories”, *Foundations of Physics* **20**, 1379–1387 (1990).
- [33] P. Heywood and M. L. G. Redhead, “Nonlocality and the kochen-specker paradox”, *Foundations of Physics* **13**, 481–499 (1983).
- [34] A. Stairs, “Quantum logic, realism, and value definiteness”, *Philosophy of Science* **50**, 578–602 (1983).
- [35] J. Walleczek, “The super-indeterminism in orthodox quantum mechanics does not implicate the reality of experimenter free will”, *Journal of Physics: Conference Series* **701**, 012005 (2016).
- [36] G. t Hooft, “The free-will postulate in quantum mechanics”, (2007).
- [37] E. Cator and K. Landsman, “Constraints on determinism: bell versus conway–kochen”, *Foundations of Physics* **44**, 10.1007/s10701-014-9815-z (2014).