**Quantum Foundations of Free Will**

# Introduction

This paper is intended to persuade an uncommitted audience that free will is illusory. I examine free will through the lens of three interpretations of quantum theory: dynamical collapse theories, hidden variable theories, and Many Worlds theories. Dynamical collapse theories, hereon called collapse theories, are the primary focus of this work since they are the most widely accepted in the current philosophy of physics climate. The core postulations and mechanics of the collapse theories are articulated. Accompanying these postulations are a few assumptions regarding the role quantum mechanics may have in one’s decisions. The postulations and assumptions together lead to the conclusion that agents are not free in the collapse theory framework. Then, I anticipate and respond to the following objections. First, that agents are at least partially free through their ability to control and change personal dispositions. Second, that the psychological level is the most appropriate scale for discussions regarding freedom. Finally, I extend my argument against free will to the hidden variables and Many Worlds theories.

## 1.1 Assumptions

I adopt the idea that deciding to do A is a momentary mental action of forming an intention to A (Mele 2017). It is that decision to A which proximately causes one to do A. What is it, though, that brings about decisions? Some philosophers argue that it is one’s own will, environment, mind, etc. However, these views tend to disregard the role that quantum mechanics may have in decision-making. The following assumptions help to show that quantum mechanics is a reasonable proximate cause of one’s decisions.

Asserting that quantum events in the brain proximately cause decisions involves a couple of assumptions centered around the mechanics of such events. Quantum events are necessarily contained in a quantum system with an associated wavefunction. Wavefunctions are time dependent mathematical entities that encode physical properties of a quantum system, such as the position and momentum of its constituents. My first assumption is that there are wavefunctions that describe the quantum events in the brain. I propose that the amplification of effects at the quantum level to the behavioral level are proximately caused by a particular kind of quantum event. Namely, a quantum event identical to the measurement of a wavefunction. The intriguing outcomes of wavefunction measurement, to which I am coming, vary among the three interpretations. My second assumption is that one’s decisions are proximately caused by a quantum event in the brain equivalent to the measurement of a wavefunction.

Actions and quantum events in the brain are connected via the measurement of a wavefunction. Measurement in the context of quantum foundations is a term that carries much baggage. Most interpretations disagree over what counts as a measurement. In general, though, measurement is characterized as an interaction between a quantum system and a classical system. One’s decision-making process involves this kind of interaction; quantum events like the movement of electrical signals and neurons (quantum system) interacting with the brain

(classical system). Hence, my assumption that quantum events proximately cause one’s decisions. Though these assumptions may seem outlandish at face value, they are relatively tame within relevant philosophical literature. For example, Robert Kane mentions quantum events of the brain in many of his works given their consistent relevance to the topic. He asserts that the brains of two identical agents with the same past may be in different states despite the sameness of their quantum mechanical description (Kane 1998). Though discussions of quantum mechanics in free will can be controversial, its presence remains. This view is meant to add to the discussion by way of palatable assumptions.

## 1.2 Scope

There are numerous variants and subvariants of the interpretations mentioned; it is beyond the scope of this paper to explain each of those differences. My aim is to highlight the core postulations of each interpretation and focus more intently on the philosophical implications that follow. In doing so, my goal is to maintain a broad yet detailed description of the interpretations that accurately relays their unique physical descriptions of the world.

## 2. Dynamical Collapse Theory

Collapse theories are borne from the Copenhagen interpretation which was pioneered and endorsed in the early 1920’s by Niels Bohr and his fellow contemporaries. The following postulations characterize collapse theories in general. First, wavefunctions prior to measurement evolve linearly and deterministically according to the Schrödinger equation. Second, wavefunctions in principle contain every measurable fact about a physical system (Maudlin 99). Lastly, a measurement of the wavefunction induces an indeterministic process that violates the smooth evolution of the wavefunction (Maudlin 99). Thus, collapse theories are characterized by a sudden violation of the deterministic evolution of the wavefunction upon measurement.

Intricacies of the collapse theory are more evident when the system is analyzed both before and after measurement occurs. Prior to the measurement of any physical observable of a quantum system, the wavefunction evolves deterministically as a linear superposition of states known as energy eigenstates. A linear superposition means that the wavefunction is in multiple states *at the same time*. Eigenstates represent every possible outcome of a measurement on the quantum system.

Each eigenstate has an associated probability referred to as prior probability. Prior probabilities reflect the probability of a particular state being realized upon measurement and are not necessarily distributed evenly among each eigenstate. The probabilities can be distributed among the eigenstates in any way, so long as the total sum of the probabilities is one. For example, consider a two-state system with state A and state B. The prior probabilities could be .50 for state A and .50 for state B. Alternatively, the prior probabilities could be .90 for state A and .10 for state B or any other mix of probabilities whose sum is one. Prior to measurement, then, the wavefunction evolves deterministically in a superposition of states, each of which having an associated prior probability of being realized upon measurement.

The system prior to measurement is straightforward, as it evolves according to well-known mechanics. However, a measurement-induced indeterministic process destroys the deterministic evolution of the wavefunction (Maudlin 99). When a physical observable of the quantum system is measured, the wavefunction collapses indeterministically to a single eigenstate of the physical observable of interest. Rather than a superposition of states, instantaneously after measurement, there remains exclusively the single eigenstate of the observable of interest. Thus, the quantum system after measurement is characterized by an indeterministic collapse of the wavefunction immediately after measurement.

Indeterminacy does not imply a lack of predictive power. Quantum theory derives its predictive power in that quantum systems can be prepared by the observer. The initial conditions of the system are at the discretion of the observer. Then, the system evolves according to the Schrödinger equation. Given that the prior probabilities of eigenstates are known, one can predict the outcomes of measurements. One can prepare numerous identical states and predict the outcomes of ensuing measurements using the prior probabilities of those states. In this way, collapse theories maintain predictive power in the probabilistic sense.

### 2.1 Decision-Making in the Collapse Theory Framework

Consider the following example of an agent in the collapse theory framework. This agent, Luz Chainge, on her daily walk to work, passes by an underprivileged individual (UI for short) asking for spare change. The outcome of this exchange goes one of three ways; Luz gives all, some, or none of her spare change to the UI. I will call the instances in which she gives at least some of her spare change the *give-state* and instances in which she does not give any change the *stinge-state*. There are other actions that might be realized in this scenario. Luz might be in the ‘food-state,’ in which she gives the UI food instead of change. Alternatively, Luz might be in the

‘pretend-state,’ in which she pretends to be too occupied on her cell phone to notice the UI. All such states are contained in the wavefunction. Even in the food-state and pretend-state, Luz either does or does not give her change. We can therefore group states such as the food-state and the pretend-state (or any alternative state) into the give-state or the stinge-state.

What causes Luz to either give or keep her spare change? Prior to her awareness of a decision, there exists a wavefunction in Luz’s brain that is in a superposition of the give-state and stinge-state. Each state has an associated prior probability. Let us assume that, given her dispositions, Luz is a notably uncharitable person. Thus, the prior probability of the give-state is significantly lower than that of the stinge-state. Assume that the prior probability of the give-state is .10, and that the prior probability is .90 for the stinge-state. Some quantum event in the brain equivalent to measurement causes the wavefunction to collapse to one of the two states.

Luz ends up in the give-state or the stinge-state depending on this indeterministic collapse.

After the collapse, Luz is said to have decided what to do with her change. Since deciding to do A proximately causes one’s doing A, we would say that Luz’s decision proximately causes her to give or not to give change. Luz gives spare change if and only if the wavefunction collapses to the give-state, in which she is aware of the decision to give her change; alternatively, Luz does not give spare change if and only if the wavefunction collapses to the stinge-state, in which she is aware of the decision to keep her change. Which action Luz performs is simply the outcome of an indeterministic collapse of the wavefunction.

## 3. Free Will in the Collapse Theory Framework

There are threats to free will lurking in this framework. There is what I call the *problem of indeterminacy*, which often arises in opposition to libertarian free will. Also, the *problem of control*, which exposes the agent’s lack of control in the collapse theory framework.

The problem of indeterminacy manifests itself directly through the mechanics of the wavefunction. Luz being in the give-state or the stinge-state, and acting accordingly by giving or keeping her change, depends on the wavefunction indeterministically collapsing to either state. Prior to collapse, the wavefunction that describes the quantum events of her decision-making process is in a superposition of the give-state (with prior probability .10) and the stinge-state (with prior probability .90). Then, a quantum event equivalent to measurement induces an indeterministic collapse to either the give-state (1/10 times on average) or the stinge-state (9/10 times on average). Luz becomes aware of this decision and acts accordingly. The problem of indeterminacy is that Luz’s actions are the result of an indeterministic collapse of the wavefunction. Luz cannot be free given that her actions are proximately caused by an indeterministic collapse of the wavefunction.

The problem of control is that Luz lacks control sufficient for free will. There is no aspect of Luz’s decision that she has control of, including the time at which a decision is made. Luz could have been aware of her decision to keep her spare change the night before she walks by the UI. Alternatively, she could have been aware of her decision the morning of her exchange with the UI. In both instances, her decision is undetermined by quantum events out of her control.

Mark Bernstein stated this sentiment nicely: one's decisions are “undetermined and result from indeterminate processes in their brain” such that Luz’s decisions are merely events that happen to her rather than something she willingly chooses (Kane 1998). In this way, agents are merely observers or vessels that carry out outcomes that are undetermined by indeterminate processes that proximately cause their decisions.

## 4. Objections

I will now respond to the following anticipated objections. The disposition objection argues that agents are free through the ability to change dispositions. The level objection attempts to show that the psychological level is the appropriate level to have discussions concerning freedom.

### 4.1 Disposition Objection

The disposition objection argues that agents are at least partially free given their ability to change dispositions. The objection calls prior probabilities into question. Why are the prior probabilities of the states what they are? One’s decisions depend largely on one’s dispositions. It is likely that one’s dispositions shape the prior probabilities of the wavefunction.

Agents frequently change dispositions, and it seems that agents are in control of this change. This sort of change is like Kane’s self-forming actions, in which an agent may exhibit dual control in which one takes responsibility for an action despite the indeterministic processes in their brain, so long as the choice was a result of the agent’s efforts (Kane 1999). Luz Chainge, for example, becomes charitable by mirroring and learning from someone who is charitable. She may then begin to do charitable things such as give her spare change to the UI more frequently. Within the collapse theory, Luz giving her spare change more frequently means that the prior probability of the give-state has increased, and the stinge-state lowered. Since Luz is in control of her change in disposition, Luz is responsible for being in the give-state more often. Therefore, Luz is at least partially free since, through a change of dispositions of which she has control, she can influence the probability of being in a particular state.

### 4.2 Response to the Disposition Objection

I agree that the prior probabilities may be shaped by one’s dispositions. This was evident in the Luz Chainge example in which I said that she is uncharitable due to her disposition. The objection claims that Luz can change her dispositions through a series of actions of which she has control. The objection examines Luz’s change in disposition too broadly. Luz can indeed become charitable through a series of actions. But how do these actions come about?

The disposition objection assumes that Luz is responsible for a change in her disposition since she is in control of her actions. It may seem as such at the macroscopic level. However, at a more minute level, Luz’s change of disposition is simply the result of a series of decisions which results in a series of actions. Luz must first make the decision to become charitable and then make the decision to pursue that desire. Luz must decide who, and in which way, she wants to mirror and learn from. Luz must finally decide to do charitable things. After that series of decisions, Luz may finally act charitably. If Luz does in fact begin to act charitably, then the prior probabilities of the give-state and stinge-state have been altered.

There is a series of decisions which results in a series of actions which describe Luz’s change in disposition. Each decision is proximately caused by unique quantum events with wavefunctions that obey collapse mechanics. Thus, each decision in the series of decisions is proximately caused by indeterministic processes out of her control. If Luz is not in control of any single decision, then she is not in control of any series of decisions. If Luz is not in control of any series of decisions, then she is not in control of any series of resulting actions. If Luz is not in control of any series of actions, then Luz cannot be in control of any change in her disposition. Therefore, the disposition objection does not hold since Luz is not responsible for any change of disposition.

### 4.3 The Level Objection

The level objection argues that discussions of freedom ought to take place at the psychological level rather than the quantum level. This objection finds traction in its appeal to mental events. Mental events such as beliefs, attitudes, and dispositions are more intelligible and more clearly linked to one’s freedom than events at the quantum level. Thus, the psychological level is the appropriate level to discuss freedom.

A convenient way to express the level objection is through Stuart T. Doyle’s example of the redness of an apple. At the atomic level, redness as a property is nonsensical since color is not a property of atoms (Doyle 2021). At the “mega-scale,” such as the entire agricultural industry, the apple is so minute that its redness is not evident (Doyle 2021). However, at the apple-scale, discussions about the apple’s redness are appropriate since its redness is clearly explained and intelligible. Analogously, Doyle claims that discussions concerning freedom should take place on the agent scale (Doyle 2021). Freedom at the quantum level is nonsensical since freedom is not a property of quantum systems. On a grand scale, such as the entire human population, the property of freedom is unintelligible. Free will is appropriately hashed out at the psychological level since the property of freedom is evident at this level via mental events. These mental events are intelligible through modern psychology and neuroscience. This line of reasoning sets out criteria that define the appropriate scale at which free will ought to be discussed.

Those apparent criteria are as follows. First, the intelligibility criterion, which requires that the events within the appropriate scale be intelligible for any productive discussion of freedom to be had. If the events on a scale are unintelligible in the context of freedom, then it is not the appropriate scale at which to discuss freedom. The relation between events and one’s freedom at that scale must be intelligible. Second, the freedom criterion, which requires that the appropriate scale contains the property of freedom.

The level objection claims that the quantum scale meets neither of the criteria. Events at the quantum scale are unintelligible; current research does not fully support there being a role for quantum mechanics in decision-making. If it were the case that quantum mechanics played a role in decision-making, it is unclear whether the outcomes of collapse mechanics are amplified to a macroscopic level. Further, the description of quantum events in relation to one’s decisions is unnecessarily complex compared to the description of events at the psychological level. Quantum events are much less intelligible than mental events. Therefore, the quantum scale does not meet the intelligibility criteria.

The quantum scale does not satisfy the freedom criteria since the property of freedom does not exist at that level. Quantum events, quantum systems, and their wavefunctions do not carry the property of freedom. Further, as stated before, it is unclear whether events at the quantum scale are realized in any macroscopic way. Analyzing freedom at the quantum level is too narrow. Quantum events do not directly lend themselves to freedom, whereas mental events do. Thus, the quantum scale does not satisfy the freedom criteria. Therefore, the quantum scale is not the appropriate level to discuss free will since the quantum scale does not meet both criteria.

The psychological level, however, satisfies both criteria. Mental events are both intelligible and intuitively linked to freedom. One’s actions are realized through mental events which the agent is, for the most part, immediately aware. One can clearly articulate their behavior in terms of their beliefs and dispositions. One cannot, however, articulate their behavior in terms of the quantum events in their brain. Thus, the psychological level satisfies the intelligibility criteria. Freedom is intuitively linked to mental events at the psychological level. Again, it is seemingly immediately obvious to agents what their beliefs and dispositions are regarding their actions. The property of freedom is most clearly seen at the psychological level. Thus, the psychological level satisfies the link to freedom criteria. Therefore, the psychological level is the appropriate level to discuss freedom.

### 4.4 Response to the Level Objection

There are two responses to this objection; first, that clarity does not necessarily imply appropriateness, and second, that quantum events are sufficient for mental events. The objection attempts to show that the psychological level is the only scale which satisfies the criteria necessary for discussing freedom. Those criteria come about given the clarity of mental events as opposed to quantum events in the context of freedom. Admittedly, the quantum events that proximately cause one’s decisions are complex. Mental events being more intelligible than quantum events does not necessarily imply that the psychological scale is the most appropriate scale. Clarity and explanatory ease are crucial in any discussion; but explanatory ease and intelligibility do not imply appropriateness.

Research in the realm of quantum foundations of free will is relatively young and difficult to flesh out. The goal of this paper is to build a foundation for such research. It will take more time to reach conclusive evidence supporting or denying quantum mechanics’ role in the philosophy of action. Until then, we cannot expect that events at the quantum level are easily explained and clearly linked to freedom. Events such as the wavefunction collapse are unsurprisingly difficult to explain in the context of decision-making. I have tried to ease the blow of this complexity by making palatable assumptions. If my assumptions hold, then events at the quantum scale certainly do link themselves to freedom (or lack thereof) via proximate causation. The complexity of the events in my view does not necessarily mean that those events do not, in fact, proximately cause one's decisions.

Rather than focusing on which level more clearly explains events, we ought to be focusing on which level contains the events that do causal work in one's decision-making. This level is where productive discussions concerning freedom ought to be hashed out. An important aspect of the level objection is its emphasis on which level contains events that proximately cause decisions. Supporters of the level objection may claim that the quantum level is, at most, only doing partial causal work alongside the causal work being done at the mental level.

Physical sufficiency provides a possible solution to the level objection. It has already been explained that mental events such as dispositions are merely a result of a series of decisions of which the agent has no control. Mental events are proximately caused by indeterministic processes in the brain. Therefore, a series of events at the quantum level are sufficient for events at the mental level. In general, then, changes at the physical level are sufficient for changes at the mental level. Jaegwon Kim expresses this line of reasoning in his brain-mind correlation thesis, which states that “the smallest change in your mental life cannot occur unless there are some specific...changes in your brain state” (Kim 1996). If physical events are sufficient for mental events, then physical events are doing some causal work in one’s decision-making. Physical sufficiency, considering Kim’s argument, shows that claims can be made for physical events doing causal work in decision-making.

Objectors may remain unconvinced by physical sufficiency given that there is still room for mental events to do causal work in one’s decision-making. This can be used as leverage to maintain that it is unclear which level does more causal work; since the psychological level is more intelligible and requires fewer assumptions, one should prefer the psychological level. However, this line of reasoning is invalid when considering another one of Kim’s ideas, known as causal overdetermination. In short, events at only one level can do causal work; there is no mix of causal work between the physical and mental level. Given that physical sufficiency supports the claim that the physical level does causal work, and that physical events are more basic than mental events, it is the physical events doing the causal work. For more on causal overdetermination, see Kim’s principle of causal exclusion.

Thus, instead of focusing on which level is more simply explained, one ought to be focused on which level contains the events doing the causal work. Given the reasoning behind physical sufficiency and causal overdetermination, the causal work in one’s decision-making is at the physical level. Therefore, the quantum level is the appropriate scale at which we ought to discuss one’s freedom.

### 5.0 Hidden Variable Theories

The mechanics of collapse theories has been described in detail in this work. The focus remained on collapse theories since they are typically taught in university physics coursework. For sake of time, I will not explain the mechanics of the hidden variable theories and many-worlds theories in much detail. I will, however, explain briefly what implications their mechanics may have on free will.

The hidden variable theories operate similarly to the collapse theories. The wavefunctions in both frameworks are similar. Prior to measurement, the wavefunction evolves deterministically according to the Schrödinger equation. The difference between the two is that hidden variable theories remain deterministic upon measurement; there is a hidden variable that determines the outcomes of measurements on a quantum system. What implications does this framework have on freedom? Let’s say Luz now operates in the hidden variable framework rather than the collapse framework. The universe and everything in it would obey deterministic mechanics. Rather than a collapse of the wavefunction, it is the relevant state of the world that determines outcomes, and therefore, proximately causes Luz’s decisions. The relevant state of the world obeys deterministic laws determined by hidden variables. Thus, the relevant state of the world is out of Luz’s control. In her own brain, too, hidden variables determine her decisions. Her decisions are the result of deterministic processes of which she has no control. One may claim that Luz can in some way alter this hidden variable by a change in disposition. It has already been shown that Luz is not in control of any change in her disposition. This means that Luz could not have any influence over an unknown variable. Thus, Luz is not free since her actions are proximately caused by hidden variables and the relevant state of the world, which are both out of her control.

Luz being in the give-state or stinge-state is a deterministic process in the hidden variable framework. Luz’s being in the give-state or the stinge-state is determined by a hidden

variable. When Luz ends up in the give-state, she was going to be in the give-state all along given that the hidden variable determined the state that the system would be in. Luz has no control over the hidden variable that proximately causes her decisions. If she is not in control of her decisions, which proximately cause her actions, then Luz is not in control of her actions. If Luz is not in control of her actions, then Luz does not have free will. Therefore, Luz is not free in the hidden variable framework.

### 5.1 Many-Worlds Theories

Many Worlds theories also obey mechanics like those of collapse theories. Prior to measurement, the system evolves deterministically according to the Schrödinger equation. Upon measurement, rather than the wavefunction collapsing to a single state, the wavefunction branches. This branching means that every state of the wavefunction is equally realized; all outcomes happen, and the wavefunction branches to accommodate each outcome. Thus, instead of a collapse of the wavefunction proximately causing one’s decisions, a branching of the wavefunction proximately causes one’s decisions.

Many Worlds theories are deterministic since every outcome is equally realized. Probability is still at play, however. There is a subjective probability which defines the agent’s probability of ending up on a particular branch of the wavefunction. Given that Luz is an uncharitable person, she is more likely to end up in the stinge-state branch. If she does in fact end up in the stinge-branch, then there exists some identical counterpart to Luz which ends up in the give-branch; identity issues lurking here will be fleshed out in a later paper. One may claim that Luz can have some control over which branch she is more likely to end up in through a change of disposition. Again, it has already been shown that Luz does not have control over any change in disposition. If Luz does not have control over any change in disposition, then Luz does not have control over which branch of the wavefunction she is more likely to be in. If Luz does not have control over which branch of the wavefunction she is more likely to be in, then Luz does not have control over her decisions. If Luz does not have control over her decisions, which proximately causes her actions, then Luz is not in control of her actions. If Luz is not in control of her actions, then Luz does not have free will in any desirable sense. Therefore, Luz is not free in the many-worlds framework.

### 6.0 Conclusion

I have attempted to persuade free will agnostics that free will is illusory. Through palatable assumptions, I have shown that quantum mechanics may play a role in one’s decision-making. My hope for this work is that it serves as the foundation for later research in the field of quantum foundations of free will. Though this work is a building block for research to come, interesting conclusions may already be drawn from the connection between quantum events and one’s freedom. I have argued that agents do not have free will in any of the three quantum interpretations. In each case, the agent’s decisions are proximately caused by either indeterministic or deterministic events in the brain. Those events, in turn, proximately cause one’s actions. Since the agent does not have control over their decisions, they do not have control over their actions. The mechanics among the three frameworks are unique, but they all deliver the same conclusion; agents do not have free will.

Word Count: 4853 [9/20/2022 - written specifically for graduate school applications]

## Works Cited

1. Mele, Alfred R. “Chapter Two.” *Aspects of Agency: Decisions, Abilities, Explanations, and Free Will*, Oxford University Press, 2017.
2. Kane, Robert. *The Significance of Free Will*. Oxford University Press, 1998.
3. Maudlin, Tim. *Philosophy of Physics: Quantum Theory*. Princeton Univ. Press, 2019.
4. Kane, Robert. “Responsibility, Luck, and Chance.” *Journal of Philosophy*, vol. 96, no. 5, 1999, pp. 217–240., [https://doi.org/10.5840/jphil199996537.](https://doi.org/10.5840/jphil199996537)
5. Doyle, Stuart T. “Sizing Up Free Will: The Scale of Compatibilism.” 2021, [https://doi.org/10.31234/osf.io/4vxqg.](https://doi.org/10.31234/osf.io/4vxqg)
6. Kim, Jaegwon. “Mind as the Brain: The Mind-Brain Identity Theory.” *Philosophy of Mind*, Westview Press, Boulder, CO, 1996.