

IN SEARCH OF FREE WILL

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF SOCIAL SCIENCES  
OF  
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

AYBÜKE KÜPCÜ YOLDAŞ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

IN

THE DEPARTMENT OF PHILOSOPHY

MARCH 2010

Approval of the Graduate School of Social Sciences.

---

Prof. Dr. Sencer Ayata  
Director

I certify that this thesis satisfies all the requirements as a thesis for the degree of Master of Science.

---

Prof. Dr. Ahmet İnam  
Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

---

Assoc. Prof. Dr. Erdiñç Sayan  
Supervisor

Examining Committee Members

Assoc. Prof. Dr. Erdiñç Sayan

---

Prof. Dr. David Grünberg

---

Assist. Prof. Dr. Hande Toffoli

---

**I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.**

Name, Last name : Aybüke, Küpcü Yoldaş

Signature:

## ABSTRACT

In Search of Free Will

KÜPCÜ YOLDAŞ, Aybüke

M.A., Department of Philosophy

Supervisor: Assoc. Prof. Dr. Erdinç Sayan

March 2010, 42 pages.

Free will is a paramount concept that is central to our everyday lives, society and moral judgements. In this thesis, I search for the conditions under which free will can exist. This is done in relation to two topics: determinism and agency. Finally, I also explore the relation between free will and several social and philosophical concepts, and discuss briefly what would be the case if there were no free will.

Keywords: free will, determinism, moral responsibility, self, consciousness

## ÖZ

Özgür İstenç Arayışı

KÜPCÜ YOLDAŞ, Aybüke

Yüksek Lisans, Felsefe Bölümü

Tez Yöneticisi: Doç. Dr. Erdinç Sayan

Mart 2010, 42 sayfa.

Özgür istenç hem günlük hayatımız hem de toplumsal hayat ve etik yargularımız açısından merkezi konumda olan önemli bir sorun. Bu tezde özgür istencin belirlenimcilik ve benlik ile olan ilişkisini inceleyip özgür istencin hangi koşullarda varolabileceğini araştırdım. Son olarak da kısaca özgür istençle bazı sosyal ve felsefi kavramlar arasındaki ilişkileri sorgulayıp özgür istenç yoksa ne olur sorusunun yanıtını aradım.

Anahtar Kelimeler: özgür istenç, determinizm, ahlaki sorumluluk, benlik, bilinç

## ACKNOWLEDGMENTS

This thesis would have not been possible without the support and contributions of numerous people to whom I would like to express my gratitudes:

Erdinç Sayan for his guidance and supervision throughout this thesis work despite its difficulties caused by my living in another country, and for his philosophy courses –*Ontology and Philosophy of Mind I – II*– which revived my interest in and belief in the importance of philosophy.

Tahir Kocayiğit for his open-mindedness and for being one of the best teachers and philosophers I have met. He taught me a lot on philosophy.

Ahmet İnam for agreeing to be my supervisor when I first started my masters long time ago. Teo Grünberg, David Grünberg and Elif Çırakman for many stimulating lectures, and Hande (Üstünel) Toffoli for her positive approach and support.

Pamela Klaassen for devoting her time to read and edit the English of this thesis.

In addition, I would like to thank my parents Şenay Küpcü and İsmail Hakkı Küpcü, and especially my brother Alptekin Küpcü, for their support. My special thanks goes to Abdullah Yoldaş – my partner, my love, my ‘yoldaş’– for always supporting me. Finally, I would like to thank myself or at least the part of me who always showed a lively interest in philosophy –despite all the prejudices and disparaging thoughts of others–, and think that thinking-questioning-discussing is a major way of searching for the reality and understanding the universe. Even when it ends up in nowhere, I think the joy of thinking is marvellous.

## TABLE OF CONTENTS

PLAGIARISM . . . . .	iii
ABSTRACT . . . . .	iv
ÖZ . . . . .	v
ACKNOWLEDGMENTS . . . . .	vi
TABLE OF CONTENTS . . . . .	vii
CHAPTER	
1 Introduction to the Problem of Free Will . . . . .	1
2 Free Will and Determinism . . . . .	4
2.1 Compatibility with a Deterministic Universe . . . . .	4
2.2 Compatibility with a Non-Deterministic Universe . . . . .	12
3 Free Will and Agency . . . . .	22
3.1 Self in Relation to Free Will . . . . .	22
3.1.1 Agent Causation Theories . . . . .	24
3.2 Free Will and Consciousness . . . . .	26
4 Empirical Constraints on Free Will . . . . .	29
5 Free Will and Everything Else . . . . .	33
6 Conclusions . . . . .	36
REFERENCES . . . . .	40

## CHAPTER 1

### Introduction to the Problem of Free Will

Having free will is an assumption we make, which constitutes the basis of our thinking and reasoning. We think our thoughts and wills belong to us, not only because they happen to occur in our brains. A thought or a will does not become my will just because it occurred in my brain, that is, it is not the physical container-ship that makes a will my will. Furthermore, although influenced by many things, we think that our thoughts and wills are not solely determined by external factors (except maybe for cases like hallucinative drug usage or brain washing).

Free will is also the basis on which our social and moral institutions are built. It is the basis of the law and the way our social system works. It is why people are punished when they do something against the laws, or why they are praised when they do the right thing according to the norms of the society they live in. It is also why we think education can change how we behave, because we believe we freely will and do things, and if we learn what is right and what is wrong (or what is useful and what is harmful), we may choose to do the right thing.

Having free will is an important attribute of us, conceived through our own eyes. In addition to this, free will is by definition important due to the power it grants to its possessor –the power to change (the future, the events, the way we are), the power to choose.

However, it is not so simple to prove the existence of free will, or that “we”, human beings, have free will.

What is free will? What do I mean by “free will”? Like many philosophical concepts, ‘free will’ does not have a clear-cut definition in the philosophical literature. According to the *The New Oxford American Dictionary*, ‘free will’ is



defined as “the power of acting without the constraint of necessity or fate; the ability to act at one’s own discretion”. According to the *Cambridge Advanced Learner’s Dictionary*, it is “the ability to decide what to do independently of any outside influence”. There is an important addition to these dictionary definitions; the concept of free will also includes the ‘freedom of will’ itself, i.e. the ability to will freely.

The external influences on free will are the easiest to distinguish and resolve. For instance, under the influence of heavy alcohol or a brain tumor, we may not behave of our own will all the time or we may not will freely, i.e. some of the things we think we will may just be the influence of the alcohol or a tumor. As science develops, we see more and more of these kind of examples of influence on will, action and decision making.

As we go to smaller and smaller scales, we start to have more and more constraints on free will. However, this may also be dangerous since we may lose track of the agent as the owner of free will when we go down to very small scales. As Dennett (2003, p. 122) once wrote, “if you make yourself really small, you can externalize virtually everything.” There is a thin line between settling the constraints of having free will and still having a meaningful agent as the owner of the free will. Nonetheless, going to micro levels in search of free will is not exaggerating or making it more difficult; it is making it more precise. The macro-level observations are probably just approximations, like the case of thermodynamics versus statistical mechanics. We would not want to stick with approximations just because it is easy and is based on easily observable properties.

In this thesis I will search for the conditions under which free will can exist and attempt to establish the existence of free will. I will mainly try to answer the following two questions:

(i) Under which conditions does free will exist?

and

(ii) Under which conditions do we (can we be said to) have free will?

This questioning will be carried out in relation to two main topics: determinism and agency.

I will also explore the relation between free will and several social and philosophical concepts, and discuss briefly what would be the case if there were no free will.

Note that, in what follows, when I say ‘universe’, I mean a physical universe with physical beings in it, including any kind of supervenient or emergent beings. Note also that I referred to agent as “it”, not as “she” or “he”, throughout this thesis, in order not to impose a gender on the agent.

## CHAPTER 2

### Free Will and Determinism

One of the main obstacles for free will seems to be determinism, although determinism as such does not possess a threat to free will. The problem can be stated by the following simple question: How can someone be free in willing or deciding if everything is pre-determined? Or more explicitly, how can we have free will, if everything we will is determined by the deterministic laws and the initial state of the universe over which we had no control or choice? This is a long discussed question that leads to the separation of philosophers into two camps on the issue: compatibilists and incompatibilists. The former are those who think that free will is compatible with determinism, and the latter are those who think that it is incompatible with determinism. So, if we rephrase the question, one of the main issues regarding free will is “Is free will compatible with determinism?” or “Can free will exist in a deterministic universe?”

To answer this question, first we need to understand what the term ‘deterministic universe’ means, and in what ways free will is compatible or incompatible with determinism.

#### 2.1 Compatibility with a Deterministic Universe

For any system to be deterministic, it needs to be a closed system governed by deterministic laws. Being a closed system means that there is no input to or output from the system. An example is a system in which energy is conserved. In physics, since matter is convertible into energy, the number of material entities need not be conserved in a system; matter can vanish but since it will vanish by being converted into another form of energy (e.g. electromagnetic or thermal),

the total energy of the system is still conserved.

For instance, consider a closed box containing oxygen molecules and ultraviolet light. The walls of the box do not interact with anything inside or outside and do not let anything including ultraviolet light in or out. The walls also isolate the box so that the temperature inside the box is not affected by the temperature outside the box. Inside the box, the oxygen molecules ( $O_2$ ) will interact with the ultraviolet light and form ozone ( $O_3$ ). Similarly the ozone will interact with the ultraviolet light and form oxygen molecules. Therefore, the number of particles at the beginning will not necessarily be equal to the number at any other time, but the energy will be conserved. Hence, since (if) all of these reactions are governed by deterministic laws, the next state of the system, i.e., the number of oxygen molecules, ozone molecules, ultraviolet photon and the thermal energy, is determined by these laws and the initial state of the system.

However, if the system is interacting with other systems, i.e., giving away particles or energy to other systems, or taking in particles or energy from other systems, it is not a closed system. Imagine we open a little window in the box full of oxygen and ozone molecules and ultraviolet light. Then the molecules and light will start to flow into the room the box is in and other type of particles, say nitrogen molecules in the air, will flow into the box. We cannot predict the state of such a system using only the laws governing the reactions in the box and the initial state of the system in the box. We need to know the laws of interaction and the number and properties of the incoming and outgoing entities. We need to include the interacting systems, i.e., the room the box is in, into our original one and form a new closed system in order to make it a deterministic system.

For a closed system, the state of the system at a given time and the (deterministic) laws that govern the system determine the state of the system at any other time. If we are living in such a system, we have no control over what is going on, even on our own thoughts or wills or decisions; basically on anything regarding ourselves (or anything else). How we are, what we feel, think, want and choose –everything is pre-determined. We are like robots in such a universe.

Actually, such a universe is itself like a robot.

Compatibilist philosophers try to find ways to define free will that is compatible with such a universe. After all, if the universe itself is a robot, there is ultimately no freedom in it, neither for us nor for anything else. Therefore I will not spend much time on the compatibilist theories, except briefly discussing the theories of two of the prominent compatibilist philosophers, Daniel Dennett (e.g. Dennett 1984, 2003) and John Martin Fischer (e.g. Fischer 1994; Fischer and Ravizza 1998).

Dennett proposes the idea that free will appears in some phase of the evolution (based on the line of thinking that evolution results in emergent properties). Dennett does not explain what “freedom” amounts to in a deterministic universe. He rather takes the evolutionary scheme and places “free will” and “consciousness” into that scheme as a property or concept originating through human evolution. Furthermore, he does not try to make free will compatible with determinism. Instead he claims that he restores responsibility –the consequence of the free will “that is worth wanting”– with his theory (Dennett 2003, pp. 162 – 245). Dennett connects freedom and responsibility to knowledge, or more correctly, he ties responsibility more to knowledge than to free will.

How would knowledge make us responsible for our wills, thoughts, feelings, and decisions in a deterministic universe? Can we be responsible for a will if we cannot will otherwise? What kind of knowledge would make us responsible in such a situation? The responsibility in question here is the responsibility of persons to the society. Therefore the knowledge at issue is related to moral values. For example, suppose we need to decide what to do in a given situation and we have two choices: the wrong thing to do, A, and the right thing to do, B, according to the moral values of the society. According to Dennett, if we know that A is wrong and B is right, then we are responsible no matter what we choose or whether or not we could have done otherwise. However, in a deterministic universe, the problem is not only that we cannot do otherwise, but we cannot choose otherwise, or will otherwise or even think otherwise. Hence, we can know

that A is wrong and B is right, but we have no say in what we will or what we choose. Therefore, we cannot choose otherwise in a deterministic universe, no matter what we know. The problem is even bigger, because what we know and what we think are also pre-determined.

What about reasoning? Even if we cannot determine what we know, we have minds and we can reason in a deterministic universe (as much as allowed by the structure of such a universe and the laws that govern it). This is also valid if we do not know the moral values of the society (or there is no moral decision on what is wrong and what is right for a given situation; i.e., it is a new type of situation that has not been encountered before). In such a case, we can try to replace knowledge with reasoning. Fischer (1994) and later Fischer and Ravizza (1998) have formulated a compatibilist theory where the moral responsibility is linked to reason-responsiveness.

Can we say we have free will if we are capable of reasoning? First of all, there are problems like how we decide the existence of this capability or how we measure it, or whether it is required that we do the actual reasoning for a decision at every decision making occasion. Nonetheless, I think granting responsibility due to reasoning is itself problematic, independent of all these issues. Consider the claim “Knowing makes us responsible” again, and replace “knowing” with “reasoning” or “being able to reason”. The claim now is that we have free will, if we are capable of reasoning (previously, if we know) no matter what we decide; against or in accordance with our reason (previously, the wrong or right thing to do). If it does not matter what we decide (against/in accordance with reason), how can it be related to the capability of reasoning? It is not different from setting a random condition for the existence of free will like saying that only those who are blond can have free will. Instead, to avoid this problem, if we say we have free will when we decide along with reason, then we eliminate our choices. How can we have free will if we cannot choose the other option (the one against reason) freely? We cannot freely choose the option which is “against reason” because by definition we did not freely will that other option. Thus, such a definition of

free will leaves us no options but one, hence takes away our freedom, instead of granting it.

Fischer and Ravizza (1998) developed a theory of moderate reasons-responsiveness, to avoid this problem of no choice that follows when we say free will exists only when we act following reason. According to them, a person is responsible for her actions if she is moderately reasons-responsive, that is, if she consistently responds to reason under certain conditions.

Consider the example of a schoolboy who delivers groceries to his old neighbour once a week. While he is playing basketball with his friends, he sees an old lady that falls down in the middle of the street and injures herself. He could quit playing and go to help the old lady, but his team is winning and he does not want to stop playing. He thinks “I probably should go help the old lady but my team is winning for the first time in a month and I really want to play”. He is responsive to reason both because he thinks he should probably be helping the old lady, and because he normally helps people who are in need. Hence, even if he decides against reason or against what he would do in a similar situation, he is still a moderately reasons-responsive person and thus morally responsible for his choice according to Fischer and Ravizza.

However, in a deterministic universe, the choice of the boy is pre-determined. In other words, even if he would behave differently –seemingly in a reasons-responsive way– in other but very similar situations, he cannot do it in this situation because he in fact has no choice or control in a deterministic universe. Everything concerning reasons-responsiveness and moral responsibility are just appearances in a deterministic universe. The behaviour of the schoolboy helping his neighbour for the groceries can seem like a reasons-responsive behaviour and a good, morally appreciated choice, but it is in fact not the boy’s response to reason nor his choice.

Harry Frankfurt’s principle of alternate possibilities (PAP) is another attempt to ensure moral responsibility within a compatibilist theory. PAP asserts that “a

person is morally responsible for what he has done only if he could have done otherwise” (Frankfurt 1969, p. 829). Frankfurt claims that incompatibilists assume the truth and necessity of PAP, but that PAP is false hence moral responsibility and determinism are compatible. Example situations aiming to show the falsity of PAP are known as Frankfurt-type examples. The generalized case can be described as follows: A person, A, wants to do something but is not sure if it is a good thing to do. Another person, B, is there at the time of the event, without being noticed by A, to make sure that the thing that A wants to do is done. Or alternatively, the natural circumstances are such that the action A is thinking of doing will occur anyway, independent of A.

For example, A is an archer and he is testing his ability to launch an arrow accurately in his backyard. He sees a bird on a tree in his backyard and thinks of shooting the bird because he feels hungry and wants to eat something. However, he is not sure if shooting the bird is really a good idea. In the meantime, B is determined to kill the same bird because it sings horribly every morning on that tree and annoys him. B knows that A does archery exercises every morning in his backyard and without being noticed B modifies the shape of A’s arrows so that they bend and shoot the bird on the tree. So even if A decides not to shoot the bird, his arrow will bend to shoot the bird and the bird will die. The argument here is that if A decides to shoot the bird and kills the bird, he is responsible for this action, even though his arrow would deviate from its trajectory and kill the bird anyhow. In other situations the forces of nature can take the part of person B. For example, the bird would have died because at time  $t$  a very sudden and strong wind occurs after A launches an arrow, causing the arrow to deviate and kill the bird. Hence, the bird would have died at  $t$  whether or not A shot to kill. Nonetheless, A would be responsible for his action if he chooses to shoot the bird.

The important point that is neglected in Frankfurt-type examples is that, the principle of alternate possibilities argument assumes the existence of free will. What it really states is that, moral responsibility exists even if we have no alternate possibilities for our actions, as long as we freely will those actions.



However, if free will is not compatible with determinism, the second part of this argument – “as long as we freely will” – is not true. Hence the existence of moral responsibility in case of no alternate possibilities (“no PAP”) is false.

Consider the same Frankfurt-type examples but this time with the lack of alternate possibilities of will instead of the alternate possibilities of action. Take the example discussed above. The bird in the backyard of A is going to die due to a sudden strong wind at time  $t$  that bends the trajectory of the arrow launched by A. This time, A does not want to kill the bird but he has an allergy to the feathers of this specific bird, which makes his brain produce certain reactions that result in him wanting to kill the bird and shuts down the logical part of the brain so that he cannot argue over his will, or his reasoning results in the idea that he is right in doing so. Hence, he could have not willed otherwise but to kill the bird. He could not have *done* otherwise but kill the bird because of the wind, and this time he also could not have *willed* otherwise but kill the bird, because of his allergy to the bird’s feathers. So even if he shoots the bird and kills it, he is not morally responsible for doing so because he could not have willed or done otherwise.

An even better example that demonstrates that freedom of will is more important for responsibility than freedom of action is as follows: Take the same situation with the bird except that the bird would not have died at  $t$  due to the wind bending the arrows from their trajectories. In this way, we eliminate the preventing forces on “could have done otherwise”. A can either kill the bird or let it live, and the bird will live if A lets it. However, A has this allergy that he is unaware of, which influences his brain in such a way that he wills to kill the bird and he cannot will anything contrary to it. He can do otherwise in the sense that there is no external restriction to it, but he cannot will otherwise. And since he cannot will otherwise or argue against his will, he will do what he wills and kill the bird. However, A cannot be morally responsible for killing the bird, when he shoots it, even if it “seems like” he could have done otherwise. In other words, A cannot be morally responsible for what he does when he cannot will otherwise.

Alternate possibilities of the will and thought are necessary conditions for moral responsibility and free will.

Fischer also proposed other examples that at first sight demonstrate a lack of alternate possibilities for will while showing the possibility of moral responsibility. The following example is from Zagzebski (2008) which she adapted from an example by Fischer (1982):

Black, an evil neurosurgeon, wishes to see White dead but is unwilling to do the deed himself. Knowing that Mary Jones also despises White and will have a single good opportunity to kill him, Black inserts a mechanism into Jones's brain that enables Black to monitor and to control Jones's neurological activity. If the activity in Jones's brain suggests that she is on the verge of deciding not to kill White when the opportunity arises, Black's mechanism will intervene and cause Jones to decide to commit the murder. On the other hand, if Jones decides to murder White on her own, the mechanism will not intervene. It will merely monitor but will not affect her neurological function. Now suppose that when the occasion arises, Jones decides to kill White without any help from Black's mechanism. In the judgment of Frankfurt and most others, Jones is morally responsible for her act. Nonetheless, it appears that she is unable to do otherwise since if she had attempted to do so, she would have been thwarted by Black's device.

The problem in this example is that it assumes that Jones's will to kill White is "free" if it is achieved without the intervention of Black. However, in a deterministic universe all wills are pre-determined, that is, "intervened." In the example above, we think that Jones is morally responsible because she decides to kill White, without the intervention of Black. That is, she is the source of her will, she is responsible for her will if *she* has the control, not Black. Free will requires not only alternate possibilities but also causal control. This is the reason for the confusion in Frankfurt-type examples including Fischer's. Causal control is indeed more important than alternate possibilities. Similar views emphasizing the importance of causal history instead of alternate possibilities have been

developed by Honderich (2002), Kane (2002) and Pereboom (2000).

What if, as suggested by the functionalist theories of mind, we are each a function? At first sight, there seems to be a similarity between the self described by functionalist theories and the self in a deterministic universe. According to functionalist theories of mind, what one does/wills/feels in a certain way can be represented by a function. If this function also determines what one does, wills and feels, than this function is who that one is; we do not get to choose it and we do not get to change it. I think the problem with a causally connected deterministic universe is even bigger than not being able to choose or change who we are. In a causally connected deterministic universe, everything depends on everything, hence we cannot isolate ourselves as functions. Hence we lack control not only over ourselves but also over the environment we live in which affects us. How does this make us responsible? We are not responsible for who we are because we lack the causal control over it, and given who we are our wills/actions are pre-determined in a deterministic Universe. We lack both causal control and alternate possibilities in a deterministic universe and therefore moral responsibility and free will are incompatible with a deterministic universe.

Since we could not find any room for free will in a deterministic universe, let us turn to indeterministic universes and see whether they are compatible with free will or not.

## 2.2 Compatibility with a Non-Deterministic Universe

There is more than one type of non-deterministic universe and not necessarily all of them are random (i.e. there are no governing laws) or governed by indeterministic laws. There can also be universes that are governed by deterministic laws but the universe itself is non-deterministic. The latter type of non-deterministic universes are especially interesting because they can provide both determinism and indeterminism depending on the conditions and situation.

If we want to keep the laws to be deterministic but let the universe be not fully deterministic, we need the universe to be an open system instead of a closed

one. This can be achieved in at least two ways:

- There is energy input into or output from the universe.
- The universe consists of causally disconnected parts.

In the first case, the state of the universe at any given time cannot fix the state of it at any other time due to the energy input/output that is not determined by the governing (deterministic) laws. Hence even if we know the laws and the state of the universe at a certain time, we cannot determine the state of the universe at any other time, unless we know the conditions, amounts and times of energy input/output (or the laws that determine these, if any). An example of this kind of open universe may be one in which emergent properties or emergent entities exist which cannot be reduced to their constituents. Nonetheless, if the emergence conditions and properties of these entities are covered by deterministic laws, then despite the energy input, the universe would still be a deterministic one. For an open universe to be indeterministic, the energy input/output should not be regulated by deterministic laws.

One such example would be a whitehole. A whitehole is a theoretical entity which is the inverse of a blackhole in astrophysics, where a star falls onto itself due to gravitational attraction and forms a singular point with such an enormous gravitational field that even light cannot escape. A whitehole is like the other end of a blackhole where the energy trapped by the blackhole comes out. In this way energy in the blackhole-whitehole system would be conserved, whereas this is not the case in an open universe. Therefore, one way to have an open universe is to have whiteholes in the universe without blackholes on the other side of them.

The idea is similar to the steady-state or quasi-steady-state theory of Hoyle and collaborators (Hoyle 1948; Bondi and Gold 1948; Hoyle, Burbidge and Narlikar 1993). According to this theory, the universe does not have a beginning like the Big Bang; instead, it is in a quasi-steady-state where matter is created “in little big bangs ... distributed over all space and time” (Hoyle, Burbidge and Narlikar 1993). The creation involves a scalar field similar to the one in

current inflation theories. This field “exerts a negative pressure ... balancing the positive energy of matter production. That permits new matter to appear in an already existing universe, instead of requiring the creation of the entire universe de novo, in a Big Bang” (Burbidge, Hoyle and Narlikar 1999). However, note that this theory conserves energy by the negative pressure created during new particle/energy creation, and the creation of new particles is constrained by laws (though these laws may be indeterministic) and conditions. This model does not seem to fit our definition of an open universe governed by deterministic laws, but it offers indeterminism in more than one way. First of all, Hoyle does not deny quantum mechanics which is indeterministic. Secondly, if the laws governing matter creation are indeterministic, this may serve as an example of an open universe with indeterministic matter input that would otherwise be deterministic (if we exclude quantum mechanics). However, note that this theory is not favoured in contemporary (astro)physics and cosmology where the Big Bang and inflation theories have a lot of supporting evidence.

Another example of an open universe starts with a Big Bang which is not a single energy explosion but an explosion that continues to power the universe by energy injection.

The second case is a universe that is made of causally disconnected parts (for which the causal disconnection can also be temporary). If a closed system has a lifetime after which the elements become independent from the system’s previous state and then form again another closed system (which is causally connected internally), then the system has choices at each time it forms a new closed system, because this new system is causally independent for an instant. This would be the case if the Big Bang/Inflation theories are correct and if the universe is expanding faster than the speed of light. In this case, parts of the universe become causally disconnected at some point. However, the current cosmological theories of this kind also require an indeterministic ingredient like quantum theory, to explain the existence of current structures that counteract the homogeneity (i.e., galaxies, stars and planets).

In the cases described above, where the universe is not fully deterministic even though the laws that govern it are, there is room for freedom provided by (i) energy input/output, or (ii) causal disconnection.

Furthermore, we can have non-deterministic universes which are governed by a set of laws that contain both deterministic and indeterministic laws, or ones in which indeterministic laws can act deterministically under certain conditions, such as the case of our universe as indicated by contemporary physics.

Assuming a non-deterministic universe, how can we have free will? Freedom as such is not enough for free will, it needs to be owned and hence controlled by the self, in order there to be “free will” (of the self). Therefore, in order to say that an agent has free will, we need the following:

1. Will should not be fully pre-determined by anything but the agent. (Other things can have an effect and a partial determination, i.e. can determine the previous wills or actions of the agent.)
2. The agent should determine the will or should contribute to the determining factors of the will.
3. The agent itself should not be fully determined by anything other than itself. This condition is necessary because if the agent is fully determined by other things, then its wills will also be fully determined by these other things, and this leads to the same problems with those of a deterministic universe.

It seems like we need determinism, for the agent should be the one determining what it wills, but we also need some kind of causal disconnection or indeterminism, to be able to break the fate that a fully deterministic universe brings and give the agent the power of control, i.e. let the agent determine, not the whole universe.

Control is the most important necessary condition of having free will. Unless what is meant by having free will is just owning it, not controlling it. However

this kind of ‘free will’ which happen to occur inside the agent hence can be said to be owned by the agent but not controlled by it, would not bring along moral responsibility. Instead, it would add some factor of randomness to lives lived in a non-deterministic universe.

Consider all kinds of non-deterministic universes to see if any can be found that can supply the necessary conditions for free will as described above.

A random universe, a universe with no laws, cannot fulfil the second condition, because nothing can be determined by anything in such a universe. Similarly, a non-deterministic universe which is governed by indeterministic laws, also cannot fulfil the second condition because the effect of an event/entity on anything is probabilistic. In such a universe an event can never be fully determined by anything, thus the first part of the second condition is not valid. What about the second part: Can an agent contribute to the determining factors in such a universe? Yes, but the contribution the agent makes can only be probabilistic/indeterministic in such a universe. Furthermore, we would still need something that discriminates the contribution of the agent as an essential one and not caused (even indeterministically) by other events.

The argument against compatibility of free will with a non-deterministic universe is as follows. If something is not a part of a causal chain (i.e. if it is not caused by something) then it happens by chance; so we depend on chance or luck for free will (see e.g. Clarke 2008). In other words, indeterministic decision making is making that decision (deciding on that particular thing) just by chance. The claim can be explained using an example involving two twin worlds, T and T', which contain the same agent in them. The agent in both worlds has to decide between two things, A and B, and the conditions leading to this decision are exactly the same in T and T'. In T, the agent decides to choose A, and in T' the agent decides to choose B. The agent causation theorists would say that the agent caused these different decisions. However, the event causation theorists would oppose to this idea saying that the two choices being different are simply due to chance or randomness. Since there is no apparent difference between

these two instances of decision making, the argument from chance/randomness is simpler; hence using Occam's razor<sup>1</sup>, most philosophers favor it.

Nonetheless, it may be said that if a non-deterministically caused event happens at the beginning of the formation of an agent, the rest (thoughts, actions and wills of the agent) depends causally mostly on the agent. Is this not better than full determinism, where nothing depend on the agent, i.e. the agent has no control over anything including itself?

Assume we live in a deterministic universe. It implies that our existence as well as our thoughts, feelings, wills and actions are determined by the initial state of the universe and the deterministic laws that govern it. At first sight, this picture seems to offer consistency. By consistency here I mean to behave, think, feel and will consistently with who we are even if we do not have any control over who we are (because it is all pre-determined). However, this turns out to be quite incorrect once we take a closer look at the deterministic universe. We are not isolated beings in such a universe; in a deterministic universe everything is determined by, hence causally connected to, the initial state of the universe together with the laws of nature. Thus, given who we are (at a certain time), we cannot determine what we will do, think, feel or will, since who we are is not the complete set of things that our actions, thoughts, feelings and wills depend on. All of these depend on and determined by the initial state of the universe and the laws of nature.

I think non-deterministic universes with causal disconnection provide us with a better picture regarding consistency. Suppose, things are deterministic until a point (in time) where a non-deterministic event occurs.<sup>2</sup> At this point the past and the future become causally disconnected. If this non-deterministic event

---

<sup>1</sup> Occam's razor tells us to choose the theory that uses fewer elements or ingredients to explain a subject matter, among many possible theories on the same subject matter. Thus, Occam's razor is a principle of economy that cuts the unnecessary parts of a theory.

<sup>2</sup> Note that indeterministic laws like those of quantum mechanics do not provide causal disconnection, since the past is connected to the future even though this connection is indeterministic to some extent.



occurs in an agent, then the actions, thoughts, feelings and wills are no longer pre-determined. Furthermore, the agent can determine its future actions, thoughts, feelings and wills, given who it is. Therefore, I conclude that in such a non-deterministic universe, we cannot fully determine who we are, but given who we are, we can determine what we do, think, feel and will. The causal disconnection in time would let us become disconnected from the past, that is, diminish the fully determining power of the past. However, this may not be enough to grant determining power to the agent alone. It seems that we need causal disconnection in space as well, i.e., the agent's actions, thoughts, feelings and wills are physically isolated/disconnected from everything other than the agent itself.<sup>3</sup> This, then leads us to the question of what an agent is. I will discuss this issue in Chapter 3. However, first the incompatibilist event-causal theories should be given a fair hearing through the most developed example available, the theory of Kane (e.g. Kane 2002, 1996).

According to Kane's theory, in order to have free will, the agent should have at least one "self-forming action/willing" (SFA) in its history. SFAs occur due to an indeterminate effort by the agent to choose among several alternative wills for each of which it has motives to will. That is, the agent indeterministically wills to do two (or more) things at the same time by showing an indeterminate effort to will/choose all of them. Hence, although the outcome is indeterministically chosen, the agent happens to choose one of the things it wants. How would this grant the agent free will? Kane's reply is that a SFA fulfils both conditions of free will: "could have willed/done otherwise" (thus, the agent faces more than one "real" choice, in the sense that what it will do/will is not pre-determined) and control (it is the agent who is choosing what it does/wills, not something else). The first condition, the existence of choices, is provided by indeterminism. According to Kane, the second condition, that choosing is under the control of the agent, is also fulfilled because the choice that is the outcome of this indeterministic

---

<sup>3</sup> One possible way to realise this spatial disconnection is by having the agent far away from everything else. The distance should be far enough to not be affected by the other things.

process is something that the agent wills since the agent had motives to will all of the input choices (Kane 2002, pp. 409–424).

One of the illustrative examples of Kane is a businesswoman who sees an assault while driving to a meeting very important for her career (Kane 2002, p. 417). She has to choose between stopping to help the assault victim or going to the meeting. She has reasons and temptation to do both of the choices, but she cannot do both of them and hence has to choose one. Here comes the “indeterminate effort” to the scene. The businesswoman makes an indeterminate effort to overcome her temptation to go to the meeting and stay to help the assault victim instead. This indeterminate effort makes this choice a SFA. The outcome of this effort, that is whether she (indeterministically) decides to stay or go, would be her choice, according to Kane, since she has reasons and motivation for both of the choices. Thus, Kane claims that a SFA satisfies both the freedom condition and the control condition necessary for free will.

A SFA is like a quantum collapse process where a superposition state collapses to one of its two (or more) possible outcomes. Kane himself states that SFAs may be chaotically amplified quantum indeterminacies occurring in the brain (Kane 2002, p. 434). Dennett objects to Kane saying that the outcome of the indeterminate choosing in a SFA may have been chosen due to some event that happened long time ago or far away (and hence the agent has no control over) (Dennett 2003, pp. 108–122). Consider the following example using the quantum analogy. The indeterminate state of businesswoman’s choosing to go or stay, collapses into the decision of going to the meeting (or staying) due to a series of events starting with a butterfly passing by and stimulating the woman’s allergy, causing her sneeze and her sneezing collapses the indeterminate state in her brain to one of two possible choices; going to the meeting. In this case a SFA does not fulfil the condition of “control”. However, I think this objection is not valid because for Kane it is not important how the indeterminate state results in that particular choice; the important part is how the possibility of more than one choice is formed, and this is due to the indeterminate effort of the agent.

The agent has the role of forming such an indeterminate state and no matter what makes it collapse into one of the choices, it will be a choice that the agent made an effort to choose. Yet, this is not the end of the story. I would argue that the same type of objection, i.e. that the agent has no control, can be made regarding the indeterminate effort of the agent. The indeterminate effort *itself* may have formed due to many other indeterminate things that the agent has no control over. Then, both forming the indeterminate effort and choosing one of the choices happens due to events that the agent has no control over. If so, such a SFA no longer satisfies the control condition for free will.

In order to overcome this problem we can question whether we really need the indeterminate effort formed by the agent or it would be enough for the agent to indeterministically choose one of the things it equally wants, e.g. by tossing a coin to decide.<sup>4</sup> Is the key point “wanting two or more things equally” or is it that “the effort is indeterminate”? I think the answer is that the effort’s being indeterminate is more important no matter how it was formed, because although the effort is indeterminate, it is still the agent’s effort. On the other hand, tossing a coin may also be regarded as a decision that the agent made, i.e., since the agent could not decide which one to choose by itself, it chooses to toss a coin to decide. Therefore, tossing a coin to choose among equally weighted wills may also be a valid method. Nonetheless, I think the important factor in Kane’s theory is not that the choice is “random” or “randomly determined” like in the case of toss-up, but that the actual process is indeterminate. However, this does not seem to overcome the control problem.

In any case, is not control of the agent over its wills an illusion? In none of the cases discussed above, the wanting of the agent is only caused by the agent; it is also caused by other things, whether the universe is deterministic or not. One way to overcome this control problem is if the agent’s contribution is necessary and is not determined by anything else, hence is indeterministic. However, I would argue

---

<sup>4</sup> This is of course just an analogy. Note that only in an indeterministic universe tossing a coin would lead to an indeterministic result.

that if in Kane's theory of indeterministic effort, the domain of indeterministic effort, that is the initial indeterminate choices, is chosen indeterministically, then again, the agent has no control over it. Kane may object to this by saying that the choice domain is formed deterministically by the agent's motives.<sup>5</sup> If so, Kane requires both determinism and indeterminism for his theory to account for free will. In this way, the agent can determine its future wills between two SFAs. Otherwise, if every will of an agent is a SFA, can it still be argued that the agent has free will or willed freely? I think not. If every will of an agent is a SFA, the conflicting wills as well as the indeterminate effort of the agent would be the results of previous indeterminate events, i.e. SFAs. This would not satisfy the second one of the necessary conditions of free will listed at the beginning of this chapter; the agent would not be the determining factor of the will or even contribute to it, since all of the agent-related factors that contribute to the will would be determined by the results of indeterminate choices.

---

<sup>5</sup> The motives themselves are formed by the agent's other properties and history, which goes back to a SFA which is indeterministically chosen and therefore breaks this regress.

## CHAPTER 3

### Free Will and Agency

What is the self that is supposed to possess free will? What does “I freely will” mean? How can I freely will? I will attempt to address these questions in two parts: first, I will examine the self in terms of conditions required to have free will, i.e. as an entity that can control what it wills, and second, I will examine free will in relation to consciousness.

#### 3.1 Self in Relation to Free Will

Recall the conditions stated in Section 2.2 that an agent needs to satisfy in order to be said to possess free will:

1. Will should not be fully pre-determined by anything other than the agent. (Anything or everything else can have an effect and a partial determination to some extent, i.e. can determine the previous wills or actions.)
2. The agent should determine the will or should contribute to the determining factors of the will.
3. The agent itself should not be fully determined by everything else.

Therefore, the agent who possess free will should have choice and control over what it chooses. As we saw already in Section 2.2, there exist no such agent in a deterministic universe, if the agent/self/mind is a purely physical entity. What if mind is not something that can be explained by laws of physics/nature? Say, for instance, it is an ontologically different level. It is a different ontological entity which does not obey to the laws of physics but perhaps the laws of mind. Would this save free will? No. As long as the laws of mind are deterministic, such a

picture will not grant free will. The underlying physical structure constrains the initial conditions of the mind, and if the laws of mind are deterministic, it is not different from the deterministic physical universe - hence the classical problem of compatibility with a deterministic universe is also valid for this case. However, if there are no laws of mind, i.e. if minds are like Gods but constrained by physical laws (but not determined by), then this picture can support/provide free will.

How is this different from dualism? It is different if minds (their properties and causal powers on physical entities) are bound by physical laws, and that they are formed by physical beings (via strong emergence), and necessitate physical beings (and hence supervene on physical). Strong emergence is defined such that “truths concerning high-level phenomenon<sup>1</sup> are not deducible even in principle from truths in low-level domain, and they are not conceptually or metaphysically necessitated.” (Chalmers 2006).

Could strong emergence satisfy the second condition or is determinism a problem for strong emergence? Can an ontological level emerge, evolve or form at some point (in time and/or when certain conditions are fulfilled) from another ontological level which is ruled by deterministic laws, and grant free will? Suppose, at the point of emergence, these two ontological levels become detached. Perhaps the lower one continues to be ruled by the same laws that used to govern it, which requires that the two levels do not interact. However, if they do not interact, the emergent ontological level has no effect on the lower level and thus it cannot provide free will. On the other hand, if they do interact, the lower ontological level is no longer ruled solely by the deterministic laws, hence also the lower level becomes indeterministic. Therefore, the only way for an emergent mind to have free will is in an indeterministic universe or by turning a deterministic universe into an indeterministic one at the point of emergence of this new level.

For the case of the emergent mind to work, i.e. account for free will and remove the permanent power of determinism, it should either

---

<sup>1</sup> *emergent phenomenon*

(i) have control over possibilities (i.e. over the possibilities provided by a probabilistic universe, such as one governed by quantum physics),

or

(ii) be an additional level onto a deterministic physical world and thus result in a total world (physical + mind) not determined by a set of laws plus initial conditions. Once minds emerge as additional ontological entities/level, such a universe is no longer deterministic.

Libet proposes two kinds of incompatibilist theories of (conscious) free will. One that effects the brain functions by non-detectable violations of physical laws, i.e. that is “at a level below that of the uncertainty allowed by quantum mechanics” (Libet 2002, p. 562). The other works by violating the physical laws while affecting the brain functions, but this time in a detectable manner, “at least in principle” (Libet 2002, p. 562). The first theory satisfies the first condition of “control over the possibilities” since the conscious will or the mind of the agent takes the control over the indeterministic process of choosing but does so in an undetectable way, i.e., the outcome choice is within the choices of the indeterministic process. The second theory proposed by Libet is in accord with the second condition of the emergence of mind that results in an indeterministic universe.

### 3.1.1 Agent Causation Theories

An important incompatibilist theory of free will is the agent-causation theory developed by several philosophers (e.g. Clarke 2002; O’Connor 2002; Taylor 1992). According to agent causation theories the free will of an agent is not caused by an event but by the agent itself. For example, a free will of turning my head towards a direction of a strange sound is not caused by my hearing the sound that triggered neurone firing in my brain which resulted in my will to turn my head towards that direction. Instead, according to agent-causal theories, I heard the sound and neurones are fired, however my will to turn my head is not causally determined by these events, but by myself as an agent. The cause of my will is not my reasons or intentions but me.

This requires a new concept of causation different from event-causation. The agent is a substance and hence causation by an agent means causation by a substance.

Agent causation is faced with the ‘ad infinitum’ argument of ‘choosing to be oneself’; what I decide depends on who I am and I cannot decide who I am because in order to decide who I am I should decide all my previous decisions and the decisions that formed them, ad infinitum (Strawson 1986, 1994). Alfred Mele replies to this argument saying “freedom is principally a feature of our actions and only derivatively of our characters from which our actions spring” (O’Connor 2002, p. 350).

“Causa sui” can be a problem for agent causation theories only if the properties of the agent, that is who the agent is, determines its actions. In this case, to be free and responsible for its actions, an agent should determine who it is, thus be causa sui. However, the deterministic relation between who the agent is and its actions/wills creates not only a problem of causa sui but also a problem of determinism. If the properties of the agent determine its actions/wills then this would not be different from an agent who lives in a deterministic universe where there is no agent causation. Therefore, for agent causation theories to avoid the same problems of compatibilist theories, they should disallow that properties of the agent determine its actions/wills. This would also solve the problem of causa sui. If the actions/wills of an agent are not determined by who it is, then the agent does not need to determine who it is, because who it is is not a determining factor anymore. Then what is the determining factor? What would distinguish agent causation from random choice?

The agent causation theory of Clarke (see e.g. Clarke 2002; O’Connor 2002) is like the quantum collapse of a superposition state, where the collapse to one of the superposed states is caused by the agent. In this picture, it seems like the agent has a power as great as the laws of nature. But, this would raise, once again, the question of what the agent is. Is it an emergent phenomenon that has the power to cause differently from event causation? How can we distinguish



between event causation and agent causation? Can an experiment be designed to measure its effects or negate it? In order for the agent causation to become a more accountable free will theory, these questions need to be addressed.

### 3.2 Free Will and Consciousness

What do we mean by self/mind in general? Is it only the conscious part of a person or the whole person? By ‘whole’ I mean including all conscious and sub-conscious processes. It is generally accepted as a personal experience when we say “I rode a bicycle today”, even though the bicycling is not a wholly consciously conducted process itself (i.e. riding a bike is a motor-like process of which we do not consciously conduct each movement). So why should ‘I’ not cover the sub-conscious processes/events, i.e. parts that conduct the subconscious processes? Indeed by using expressions like “I am riding a bike,” in folk psychology we already accept that ‘self’ also covers subconscious processes. Defining self as only the conscious part of a person makes it a fighting/competing entity with the subconscious part of that person. It would not be different from multi-personality disorder cases in psychiatry.

In the 1980s, Benjamin Libet conducted an experiment measuring brain activity during conscious decisions. The subjects were asked to move their hand at an arbitrary moment decided by them, and to report when they made the decision (they timed the decision by noticing the position of a dot circling a clock face) (Libet et al. 1983). The electrical activity of their brain was monitored during this process. As a result of the experiment, Libet found that the reported times of the decisions were  $\sim 350$  ms after the so-called readiness potential which marks the beginning of the movement (Libet et al. 1983). This much debated finding is taken by Libet to show that the conscious decisions are determined unconsciously beforehand but there is enough time to consciously “veto” them (Libet 2002) .

I think Libet’s experiments only show that consciousness takes time, that is for anything to be consciously perceived by the self there is a minimum duration necessary. This is most probably related to the nature of consciousness,

in the sense that a state is conscious when there is brain activity almost everywhere in the brain. Recent experiments by Gaillard et al. (2009) showed that conscious states are different from unconscious states by having a sustained electrical activity longer than 300 ms “across a great variety of cortical sites, with a special involvement of the prefrontal lobes....suggesting conscious perception is broadcasted widely across the cortex.”. This suggests that the information is distributed to every section of the brain for a state to become a conscious state. This distribution of information throughout the brain takes a certain time, and is the reason for the delay between the start of the brain activity and the report of “freely willed decision of movement” by the subjects of Libet’s experiment. The delay is due to the time necessary to distribute the information of the “freely willed decision of movement” to be consciously perceived by the subjects, not the actual “freely willed decision of movement” to take place.

Another interpretation of the findings in Libet’s experiments is that free will does not always have to occur at the conscious level. It seems reasonable if we think of the number of actions that take place in the human body and brain at the subconscious level. It may be that basic wills and decisions do not require the agent to be conscious during the whole process, but more complex wills and decisions do. The “free will experiments” like Libet’s, measure very basic acts/wills that do not require consciousness. Thus they may be examples of un/subconscious free wills. There is yet no experiment that shows the non-existence of conscious free wills in case of complex decisions, because we do not yet have the knowledge to disentangle the signals of the various processes for such a complex case from the “signal” of free will.

I think what Libet’s and other recent neurological experiments, such as those of Gaillard et al. (2009), show is that consciousness is not a level which exists permanently and acts as a container. Our thoughts, wills and feelings do not stay, without being noticed, at the subconscious level and then transfer themselves from the subconscious to conscious level. There is no such thing as a conscious level that hosts multiple states like toothache, sight of a colorful autumn leaf on

the street and happiness. Rather, these neurological experiments show that there are conscious states, like toothache, that form in time and then dissolve in time.

We relate free will with consciousness because we think we can control what we do, will or decide when we are conscious. However, this is only an illusion in a deterministic universe. In the case of an indeterministic universe, consciousness can only have the control that is provided by certain types of indeterministic universes, if any, as discussed in Section 2.2. Otherwise, free will requires some kind of agent causation or emergent non-deterministic mind. In order to determine how consciousness can provide an agent the control over its wills, in an indeterministic universe or for agent causation theories, we need to understand what consciousness is which is out of the scope of this thesis.

## CHAPTER 4

### Empirical Constraints on Free Will

We have some viable theories of free will like that of Kane, agent causation theories, and the mind as an emergent entity (see Sections 2.2, 3.1.1 and 3.1 for details, respectively). Is there a way to empirically distinguish these theories? Or is there a way to detect the existence of free will, i.e. establish some empirical conditions that a free will needs to satisfy?

Kane's theory, agent causation theories and the theory of emergent non-deterministic mind require that free will is not a determined outcome of deterministic natural laws. What about indeterministic natural laws: Would the will/choice of the agent agree with the probabilistic outcomes of indeterministic choices? According to Kane's theory of free will the answer is yes<sup>1</sup>. However, the agent causation theories and the emergent non-deterministic mind theory do not explicitly mention anything about whether the choices of the agent would comply with a universe governed by indeterministic laws.

For example, according to the indeterministic laws an agent has a 60% probability to will an ice-cream and 40% probability to will a chocolate bar. Would the will of the agent be either ice-cream or a chocolate bar, or would it be something completely different, say, an apple, according to the agent causation theories and the emergent non-deterministic mind theory? Or would it be one of the two indeterministically constrained choices but with different probabilities?

Let us examine all three situations where:

1. The agent wills in accord with the indeterministically constrained choices and their probabilities.

---

<sup>1</sup> This positive answer results in objections to Kane's theory that the agent lacks the control over her wills/choices.

2. The agent wills in accord with indeterministically constrained choices but with different probabilities.<sup>2</sup>
3. The agent wills something different than the indeterministically constrained choices.

The first case is the one that complies with Kane's theory of free will and possibly with some of the agent causation theories. The problem is that it gives rise to objections that the agent has no control over its wills/choices. That is, if wills are in accord with predictions of indeterministic laws, then it is the indeterministic laws that "determine" those wills, not the agent. Kane tries to overcome this by saying that the agent has motives to choose either of the choices, and since it is the result of the agents's indeterministic effort, it is the agent's choice (see Sec. 2.2 for further discussion). The agent causation theorists may object to this, arguing that the agent is the one who determines which of the indeterministically constrained choices will be chosen at a given time. This claim is similar to the non-deterministic free will theories mentioned by Libet where the effect of the agent is smaller than the uncertainty allowed by quantum mechanics, which makes the effect non-detectable but there<sup>3</sup>. However, since this is a non-detectable effect both in reality and in principle, agent causation theories that claim this remain metaphysical theories<sup>4</sup> cut by Occam's razor unless any other evidence or ways to falsify it are provided. Furthermore, the indeterministic outcomes are observed in cases which do not involve any agent. Therefore, the effect of the existence of the agent in a decision moment can be said to be redundant.

The second and third cases seem to overcome the problem of non-detectability

---

<sup>2</sup> This specific case is suggested by Erdinç Sayan (private communication).

<sup>3</sup> At first sight it also looks similar to the hidden variables theory of quantum mechanics where in this case something of the agent is a hidden variable that constrains which of the choices will be chosen. However, hidden variables theory is a deterministic theory, suggesting that if it is true there are no choices and everything is pre-determined. Thus, it is not a viable option to establish free will.

<sup>4</sup> As metaphysical theories I think the explanation of agent causation theories are at the same level as explanations from natural laws. That the agent chooses which possibility is chosen at a given time is metaphysically the same as to say that which possibility is chosen at the time is indeterministically determined by the natural laws (and there is no specific explanation on why one of those possibilities is chosen, rather than another, at a given time).

and indistinguishability. They argue that the agent's wills do not obey natural laws, whether deterministic or indeterministic, hence the agent has control over its wills rather than nature, and that this effect is detectable. This effect can be detected such that every time an agent freely wills something, we should see a law violation<sup>5</sup> in the agent's physical being (i.e., in the brain of human beings).

Suppose we see a law violation in a person's brain. Would we immediately say that the agent's current will/act/thought is a free will/act/thought, or would we look for any additional constraint like reason-responsiveness? Assume the agent willed something that appears totally crazy and inconsistent with the agent's personality, and at the same time we detected a law violation. Is this still a free will or not? Would we say no because it is not a reasons-responsive will, that it does not seem to be in accord with any of the rational reasons that the agent has at that time for such a will? I think this is a dangerous road to take.

Placing reasons-responsiveness as a constraint works well against the problem of randomness. The problem of randomness is how to distinguish a will that is "freely willed by an agent" from a random will. If the will does not comply with the character of the agent, and its motives, why think that it is freely willed by the agent rather than a random event that happened to the agent? It seems an additional constraint is needed; either reasons-responsiveness or a (micro-level) pattern of law violation that is not random.

If we put the additional constraint that the free wills of an agent are those which violate laws and have rational outcomes for the agent, then would not it be the same as saying that those wills obey a type of 'rationality law' and are hence controlled by it rather than the agent? If this is the case, we could as well accept any reasons-responsive compatibilist theory of free will since according to those theories the wills of the agent are also controlled by some kind of rationality law, only this time the laws it is based on are deterministic.

---

<sup>5</sup> The law violations should be such that they cannot be explained by changing the laws of nature in a consistent way. That is even if the physical/chemical/biological laws are modified in the future, it should always be the case that they cannot explain the rest of the nature and these events at the same time in consistent way.

What about a law violation pattern that is not random, i.e., more like the second case (that the agent wills in the domain of indeterministically constrained possibilities but with different probabilities) rather than the third case (that the agent wills something completely different)? Note that such a pattern must not be consistent in macro-scale, that is the pattern an agent shows for one type of its wills should not be similar to the pattern the agent shows for another type of its wills, and there should not be a connection between these micro patterns and the properties of the wills. This condition is necessary to ensure that there exists no law formulation that can include these micro patterns into laws of nature resulting in the willing of the agent obeying the laws of nature whether deterministic or not. Thus there should be a micro-pattern for each free will of the agent but no macro-pattern.

Consider this problem using the previous example that the agent can will an ice-cream with 60% probability or a chocolate bar with 40% probability according to indeterministic laws. The micro-pattern theory requires that given the same situation, the agent should will an ice-cream for 30% of the cases or a chocolate bar for 70% of the cases. This is not random because the realisation ratio of different choices is not 50–50, and the agent does not will something completely different like an apple. Furthermore, it is not fully constrained by indeterministic laws because the probabilities of the choices are different from 60% vs. 40%, and they do not obey any other laws on the macro-scale. Given such a case, would we say that the agent freely willed something even if it does not comply with the character and motives of the agent? I think yes, we would.

The key element of this scheme is provided by consciousness or some other property of mind which overcomes the indeterminacy – or more correctly, which can rule inside the uncertainty margins of the indeterministic laws of nature. This property of mind should be an emergent property so that it can escape from the domain of natural laws and only found in agents.

## CHAPTER 5

### Free Will and Everything Else

Free will is important not only because it is the basis of moral responsibility and desert, but it is also the basis of how we think about the world and ourselves.

How would a free will-less universe affect us, change our way of thinking and feeling, or change the meanings we attribute to many important concepts? It seems that the effect is not foreseeable –unless the universe is deterministic and we have tools to predict the next state of the universe– and we have no control over it, if there is no free will. Even though judging people based on the idea that they are morally responsible is wrong in a free will-less universe, there is also nothing we can really do about it. Every thought, feeling, will and action of us will be a deterministic or indeterministic outcome of the natural laws. We can of course try to do what we “believe” is the “right” thing, but whether we succeed or not, whether the outcome is what we “aimed” for or not, does not depend on us. Therefore actually, even if we find-out that there is no free will, the effects of this thought will occur independent of us since how we think, feel, will, act is not controlled by us in such a universe.

I think one important problem of a free will-less universe is whether one can trust any thought. This is not straightforward to answer, because in a free will-less universe, thoughts are not generated or controlled by the agent itself but as everything else the thoughts are determined (partially in cases of indeterministic universes) by the governing laws and the initial state of the universe<sup>1</sup>. Therefore, even though an agent may think that it is following a line of logic, a line of reason while thinking, every thought the agent has is in fact determined (partially in

---

<sup>1</sup> In case of an open universe, another determining factor is what provides the openness of the universe, i.e. energy input/output



cases of indeterministic universes) by the laws (if any) that govern the universe and the initial state of the universe. Following logic or reasoning are in fact following the natural laws governing the universe. How natural laws affect our way of thinking macroscopically can be seen today. The widely used two-value logic that governs our truth value judgements, is influenced by the macroscopic physics which is observed in and rule (even though it is an approximation) our everyday lives. Alternatively, quantum logic offers a different path of rationalizing and truth values. Are we indeed following quantum logic because quantum physics is what governs our brains and the universe? If so, this may be the reason why some people sometimes contradict the main stream two-value rationality and moral judgements.

All in all, if there is no free will, independent of whether the universe is fully deterministic, indeterministic or completely random, for all the things we do, think and feel, we follow the logic (or randomness) that governs the universe we live in. Our rationality and moral values, not only the moral values of one person, but the concept of “rationality” and generally accepted moral values and judgements, would just be the outcome of how the universe is ruled by natural laws.

Furthermore, in a causally connected universe, whether deterministic or not, rationality is an approximation of the logic behind governing laws because in a causally connected universe, everything depends on everything else and hence what happens now and here, and how we evaluate it, does not only depend on natural laws but also on the conditions and state of the universe. Hence in such a universe rationality may not even be something consistent in itself.

In a universe without free will, the process of measurement and evaluation cannot be isolated from what is measured. In fact, the observed phenomenon, the experiment that measures it and the agents that evaluate the results are all ruled by the same laws. Therefore, the process of measurement and evaluation should be considered as a part of the experiment. There is a similar construct in quantum physics, where the measurement process couples to the measured phenomenon

and actually changes the measured phenomenon. In a universe without free will, the evaluation process should be added on top of these two. The evaluation process is also coupled to the measurement process and the measured phenomenon in the sense that all are ruled by the same laws. Hence, the evaluation process must be included in the evaluation.

## CHAPTER 6

### Conclusions

The existence of free will has one major condition: control. For there to be free will, the agent –not someone or something else– has to have control over its will. Thus the agent should determine its wills. This does not mean that the agent has to be the only determining source of the wills, but the agent’s contribution should be an essential one that is not determined by anything or anyone other than the agent. The condition of control also requires that the contribution of the agent should not be determined by any set of deterministic laws, because in such a case it is natural laws that determine the wills of the agent, not the agent itself.

The issue of control is the main obstacle against compatibilistic (deterministic) free will theories since a deterministic universe rules out the existence of free will since the agent in such a universe can never determine/control (be the determining/controlling source of) its wills.

What happens if the laws are indeterministic? The issue of control is actually a major obstacle not only for compatibilistic (deterministic) free will theories but also for indeterministic free will theories. Indeterminism provides some kind of freedom (that there is not a single outcome of the past events and hence the future is not fully pre-determined by the past) in the universe but that freedom is not enough to assure the existence of free will. Indeterminism provides freedom but does not seem to give control of this freedom to the agent. One of the most advanced indeterministic event-causal theories of free will, the theory of Robert Kane, tries to overcome this control problem by saying that the will that was indeterministically chosen is still the will of the agent because the agent had motives to will it and actually showed an indeterminate effort for it. However,

there is still the control problem regarding who determines the domain of these indeterminate choices.

It seems that for the agent to have control of its wills either it has control over the possibilities of an indeterministic universe, or that the mind (or the controlling part) of the agent is an emergent entity or property which can affect the rest of the universe but is not affected by it and does not obey any laws whether deterministic or indeterministic.

Another way to overcome the control problem is suggested by agent causation theories. According to agent causation theories the free will of an agent is not caused by an event but by the agent itself. This seems to overcome the problems of determination and control, however agent causation theories have to account for the questions introduced by this new concept of 'causation', like: "What is the difference between event causation and agent causation? How can its effects be measured? Why do only agents have this ability to cause but not all substances?". Furthermore, the role of consciousness in this type of causation, or even in general with regard to free will, is still not well understood.

Given these conditions on free will, we may look for empirical evidence for its existence. First of all, we need indeterminism. However, if the universe is governed by indeterministic laws, how can we distinguish between the outcome of these laws and the effect of the agent on the free will? The problem with free will theories that claim that the agent affects the choices within the levels of uncertainties, is that there is no way to distinguish, measure, or prove this effect. Furthermore, it seems redundant to attribute an agent such a power given that indeterministic results are also obtained for situations that do not involve an agent.

Instead, we may look for a more obvious and measurable effect as evidence for free will, namely that the agent should change the probabilities of the outcomes of indeterministic processes, and it should do so without obeying any other kind of law. We also need some constraints that would ensure that this effect is the effect of the agent and not something random. One such constraint would be

that this indeterministic process with changed probability should occur in the agent's body. Another may be a kind of reasons-responsiveness: that the agent responds to reason and its wills are usually rational and in compliance with its character. However, such a constraint may also impose that there is a macro scale 'rationality law' that determines an agent's wills. It is this rationality law that changes the probabilities of the indeterministic processes, not the agent. Thus, such a reasons-responsiveness constraint would in fact take the control away from the agent instead of endowing the agent with it.

In conclusion, we are left with four constraints on the possible empirical evidence for free will:

- 1) The universe should be governed by indeterministic laws.
- 2) Probabilities of the indeterministic outcomes of the process that leads to a will, should be different than the probabilities that are imposed by the indeterministic laws.
- 3) These changes in the probabilities of the indeterministic outcomes should not obey any law, neither deterministic nor indeterministic.
- 4) This indeterministic process that leads to a will, with probabilities different than those imposed by indeterministic laws, should occur in the body of the agent.

These are necessary conditions for free will but they may not be sufficient for the existence of free will.

What if we do not observe any empirical evidence for free will? What if we are living in a deterministic universe where there is no room for free will? Free will is not only the basis of the concepts of moral responsibility and reward, but also the basis of how we think about the world and about ourselves. Unfortunately it seems that we can do nothing about it, if there is no free will. If there is no free will, we have no power to change or effect how things are or will be, thus, even though judging people based on the idea that they are morally responsible is wrong if there is no free will, there is also nothing we can really do about it, since every thought, feeling, will and action will be a deterministic or indeterministic outcome of natural laws. We can of course try to do what we "believe" is the

“right” thing, but whether we succeed or not, that is whether the outcome is what we “aimed” for or not, does not depend on us <sup>1</sup>.

---

<sup>1</sup> Personally, I would prefer to live in an indeterministic universe even if there is no free will. It will at least be the case that nothing is strictly pre-determined – “If I cannot determine them, nothing should.” And some things will come as a surprise – “Let’s see what will happen to me today.”

## REFERENCES

- Bondi H. and T. Gold, “The Steady-State Theory of the Expanding Universe”, *Monthly Notices of the Royal Astronomical Society* 108 (1948), 252
- Burbidge, G., Hoyle, F. and Narlikar, J. V., “A Different Approach to Cosmology”, *Physics Today* 52 (1999), 38–44
- Cambridge Advanced Learners Dictionary*, Third Edition (Cambridge, UK: Cambridge University Press, 2008) <http://dictionary.cambridge.org/>
- Chalmers, D., “Strong and Weak Emergence” in P. Davies & P. Clayton (eds.), *The Re-Emergence of Emergence* (Oxford: Oxford University Press, 2006)
- Clarke, R., “Libertarian Views: Critical Survey of Noncausal and Event-Causal Accounts of Free Agency” in Robert Kane (ed.) *The Oxford Handbook of Free Will* (New York: Oxford University Press, 2002)
- Clarke, R., “Incompatibilist (Nondeterministic) Theories of Free Will”, *The Stanford Encyclopedia of Philosophy* (September 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/entries/incompatibilism-theories/>
- Dennett, D. C., *Freedom Evolves* (New York: Viking Penguin, 2003)
- Dennett, D. C., *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, MA: MIT Press and Oxford University Press, 1984)
- Fischer, J. M., “Responsibility and Control” *Journal of Philosophy* 79 (1982), 24–40
- Fischer, J. M., *The Metaphysics of Free Will* (Oxford: Blackwell, 1994)
- Fischer, J. M. and Ravizza, M., *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge: Cambridge University Press, 1998)

- Frankfurt, H., “Alternate Possibilities and Moral Responsibility” *Journal of Philosophy* 66 (1969), 829-39
- Gaillard R. et al., “Converging Intracranial Markers of Conscious Access”, *Public Library of Science Biology* 7(3) (2009), 1–21
- Honderich, T., “Determinism as True, Compatibilism and Incompatibilism as Both False, and the Real Problem” in Robert Kane (ed.) *The Oxford Handbook of Free Will* (New York: Oxford University Press, 2002)
- Hoyle, F., “A New Model for the Expanding Universe”, *Monthly Notices of the Royal Astronomical Society* 108 (1948), 372-383
- Hoyle, F., Burbidge G., and Narlikar, J. V. , “A Quasi-Steady State Cosmological Model with Creation of Matter”, *The Astrophysical Journal* 410 (1993), 437–457
- Jewell, E. J. and Abate, F. R. (eds.), *The New Oxford American Dictionary* (Oxford Oxfordshire: Oxford University Press, 2001)
- Kane, R., “Some Neglected Pathways in the Free Will Labyrinth” in Robert Kane (ed.) *The Oxford Handbook of Free Will* (New York: Oxford University Press, 2002)
- Kane, R., *The Significance of Free Will* (New York: Oxford University Press, 1996)
- Libet, B., Gleason, C. A., Wright, E. W., Pearl, D. K., “Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act”, *Brain* 106(3) (1983), 623-642
- Libet, B., “Do We Have Free Will?” in Robert Kane (ed.) *The Oxford Handbook of Free Will* (New York: Oxford University Press, 2002)



- Pereboom, D., “Alternate Possibilities and Causal Histories” *Philosophical Perspectives* 20(2000), ed. James Tomberlin
- O’Connor, T. “Libertarian Views: Dualist and Agent-Causal Theories” in Robert Kane (ed.) *The Oxford Handbook of Free Will* (New York: Oxford University Press, 2002)
- Strawson, G., *Freedom and Belief* (Oxford: Oxford University Press, 1986)
- Strawson, G., “The Impossibility of Moral Responsibility” *Philosophical Studies* 75 (1994), 5–24
- Taylor, T., *Metaphysics* (Englewood Cliffs: Prentice-Hall, 1992)
- Zagzebski, L., “Foreknowledge and Free Will”, *The Stanford Encyclopedia of Philosophy* (March 2008 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/entries/free-will-foreknowledge/>