

Free Will and Moral Responsibility

Bachelor Thesis Psychological Research Methods

Paul Lodder

University of Amsterdam, Amsterdam, The Netherlands

For centuries, philosophers have been pondering the problems of whether human beings have free will and whether they can still be held morally responsible for their actions if it turns out that they do not have free will. Recently, there have also been conducted a number of scientific studies on these topics. The present thesis uses those recently conducted experiments in order to formulate an answer to those two age-old problems relating to free will. It turns out there are good reasons to assume that free will is illusory and that our experience of having free will is constructed by our brains. Whether people can - given the previous conclusion - still be held morally responsible for their actions, remains unclear after reviewing the recently conducted experiments. Given the illusory nature of free will, it is argued that people are not morally responsible in an ontological sense. Yet, it is argued that they can still be held morally responsible as based on legal conventions.

We must believe in free will - we have no choice
(Singer, 1997)

Introduction

The paradoxical nature of the quote above portrays the profound complexity of the problem of free will. This conundrum has plagued philosophers for centuries and can be traced back all the way to the era of the pre-Socratic philosophers. A few decades ago, neuroscientists entered the debate and claimed that we could possibly solve the problem by studying the brain. Roughly, the problem of free will boils down to two basic questions: (1) Do we as human beings have free will? and (2) Are we morally responsible for our actions? In the present study I will try to answer both of these questions by using philosophical reasoning and scientific evidence.

Scrutiny of the literature about the problem of free will showed that key concepts are ambiguous. Performing a thorough analysis will only become possible after first giving a clear definition of free will and the other concepts used in this thesis. Therefore, I will first clarify the key concepts and subsequently indicate which definitions I will use in my analysis.

First, it is important to provide the definition of *determinism*, a concept that is often contrasted with free will (Smilansky, 2000; Strawson, 1986) and therefore central to the topic of the present article. The definition of determinism that I will use states that at any instant exactly one future is

compatible with the state of the universe at that instant and the laws of nature (Mele, 2006, p. 3). I use this definition because it adequately captures the way in which determinism is used in the articles describe in this thesis. The definition implies that knowing the present state of the universe combined with the laws of nature enables us to predict what the future state of the universe will be. It is important to distinguish determinism from fatalism, because people without a background in philosophy tend to confuse the two concepts now and then (Feltz, Cokely & Nadelhoffer, 2009). *Fatalism* means that some or all events in the universe are destined to happen, independent of causal influences that precede the event. If fatalism is true, then a particular event will happen at some point in the future and there is nothing that can be done or occur to prevent it from happening.

Another concept in need of clarification is free will itself, which is a little bit more difficult. In fact, I will propose three different definitions of free will and subsequently indicate which definition is of interest for the present thesis. First, there is the kind of free will that nearly everyone is familiar with. It is the subjective experience of making a choice of which you feel you are the agent that makes it. From now on, I will call this type of free will *phenomenological freedom*. The second kind of free will can be seen as a freedom in acting without being constrained by external factors, such as legislation or coercion. In other words: the ability to act in accordance with what you want to do. I will dub this kind of free will *behavioral freedom*. The last type of free will is the one I will focus on in the present article. It is free will in the most fundamental and deepest sense and I therefore call it *metaphysical freedom*. It is the kind of freedom that people would have if – given the prior state of the universe combined with the laws of nature – they could still make a decision that was not the result of the preceding deterministic causal mechanism.

While debating about determinism and free will, it has to be the metaphysical freedom that is related to determinism.

I would like to thank Dingmar van Eck for supervising my bachelor's thesis and providing me with important insights about the nature of free will and moral responsibility.

► E-mail: p.lodder@uva.nl

The other two types of freedom are perfectly compatible with determinism. For example, a person can have the subjective experience of making the choice to buy a cup of coffee (phenomenological freedom) without being constrained by external factors such as closed coffee shops or being out of money (behavioral freedom); but at the same time it is still possible that the decision to buy a cup of coffee was purely determined by the prior state of the world combined with the laws of nature. Some philosophers, however, argue that given the prior state of the universe combined with laws of nature – in other words: given the truth of determinism – it is still possible for the person determined to buy coffee to suddenly *not* buy the cup of coffee. This person would then have a metaphysical kind of freedom and the question of whether this metaphysical freedom is compatible with determinism remains one of most complex philosophical problems of our time.

Within this debate there are in general three different positions. If you believe that determinism is true but that it's at the same time no problem to have metaphysical freedom, then you are a *compatibilist* (Dennett, 2003). If, on the other hand, you think that determinism and metaphysical freedom are not compatible then you are an *incompatibilist*. There are, however, two kinds of incompatibilists. The *hard determinists* (Smilansky, 2000; Strawson, 1986) claim that determinism is true and that it is incompatible with metaphysical freedom and that the latter therefore does not exist. *Libertarianists* (van Inwagen, 1983; Kane, 1999) argue that metaphysical freedom and determinism are incompatible and that human beings are free agents, hence, determinism must be false.

A final concept in need of clarification is moral responsibility. First, moral responsibility will have to be distinguished from two other definitions of responsibility, namely causal responsibility and legal responsibility (Miller & Feltz, 2009). *Causal responsibility* is a purely descriptive way of saying that something is the direct cause of something else. For example, the rain is the cause of the flood. The difference with moral responsibility is that causal responsibility is purely descriptive and therefore lacks a normative component through which a particular event or act is evaluated. *Legal responsibility* can be defined as being held responsible for an act based on a judgment within our legal system. The difference with moral responsibility is that legal responsibility does not have to be a correct assessment of what actually happened because it is based upon acquired evidence that does not always match reality. One could for example be held legally responsible for a murder while actually being innocent (Miller & Feltz, 2009). Given the differences with the other two kind of responsibility, *moral responsibility* can be defined as the normative evaluation of a person that committed a particular act. If the act is seen as positive, then the person can be evaluated as praiseworthy, while if the act is negative, the person can be seen as blameworthy. Another contemporary debate within philosophy is whether people can be held morally responsible if it turns out that they do not have metaphysical freedom (Frankfurt, 1969).

The aims of the present article are twofold. First, I will investigate whether there are plausible grounds for the perspective that people do not have metaphysical freedom. Indeed, I will conclude that it is plausible to assume that people do not have this kind of freedom. Given this conclusion, I also investigate the question whether people can still be held morally responsible if they do not have metaphysical freedom.

The outline of this thesis is as follows. Chapter one focuses on the classic neuroscientific experiments conducted by Libet, including the most important criticisms. Subsequently, in chapter two the recent scientific studies about the existence of metaphysical freedom will be elaborated and I will argue that this kind of free will is illusory. In chapter three I will give an overview of the recent experiments on the topic of free will and moral responsibility. Finally, in the discussion I will perform a conceptual analysis and try to integrate the findings of the described scientific experiments. I will argue that both metaphysical freedom and moral responsibility are incompatible with determinism. Yet, people can still be held morally responsible as based on legal conventions.

Libet

Benjamin Libet can be considered the founder of scientific research into the topic of the existence of metaphysical freedom. He was the first to actually conduct a controlled scientific study in order to find out whether present day neuroscience leaves any room for metaphysical freedom (Libet et al., 1983). His research is based upon a study by Kornhuber and Deecke (1965) who showed that the decision of people to flex their fingers is preceded by electrical activity in a part of the motor cortex called the *supplementary motor area* (SMA). This occurrence of electrical activity approximately 200-1500 milliseconds before the actual movement was coined the *readiness potential* (RP).

This finding raises an interesting question related to the origin of our actions. Does the conscious experience of deciding to move happen *before* or *after* that brain activity? If brain activity related to the decision to move precedes the conscious awareness of the decision, then this seems to be an indication that the movement was initiated by unconscious brain processes, and not by our own metaphysical freedom. Yet, if conscious awareness precedes the RP then this would leave open the possibility of having metaphysical freedom. This paragraph will give an overview of the studies conducted by Libet in which he addressed this issue.

Libet's famous experiments

First, it is important to make a distinction between two kinds of intentionality: *prior intention* and *intention-in-action* (Searle, 1983). Prior intentions precede an action by a considerable amount of time. The plan to get married next year can be seen as an example of a prior intention. An intention-in-action, on the other hand, is just in advance of a particular action. An example is suddenly hitting someone

while getting emotional during an argument. The decisions participants have to make in the following neuroscientific experiments are about intention-in-action and not about prior intentions.

The validity of the readiness potential (RP) as an operationalization of the brain activity preceding a motor act has been confirmed (Libet, Wright and Gleason, 1982). In order to find out whether brain activity related to the decision to move in reality precedes the conscious awareness of that decision, Libet and his colleagues (1983) asked five participants to watch a rotating spotlight that mimicked the movement of the arrow of a clock. Subsequently, participants were asked to flex their fingers at a moment of their own choice. After they made the movement participants had to indicate the position of the spotlight at the time they first became conscious of deciding to make the movement. The brain activity in the SMA of each participant was measured with electroencephalography (EEG) and the moment they flexed their fingers was recorded with electromyography (EMG). The results indicated that on average the brain activity *preceded* the movement with 550 milliseconds and the awareness of deciding to move with 350 milliseconds. Libet and his colleagues concluded that the initiation of a voluntary act begins unconsciously and the subjective awareness of the decision to act happens after the brain activity that is related to that decision.

While some would conclude that this experiment indicates the non-existence of our metaphysical freedom, Libet himself does not draw this conclusion. He states that it is still possible to have metaphysical freedom by means of *vetoing* an action during the 150 to 200 millisecond interval between the act and the conscious awareness of the decision to act. He argues that during this interval, the onset of an unconsciously initiated movement can be blocked by an act of metaphysical freedom. This hypothesis is supported by another study conducted in Libet's lab (1985). He asked five participants to flex their fingers at pre-arranged times, but during half of the trials, they were asked to veto their movement as soon as they become aware of the decision to move. It turned out that participants were indeed able to exert a veto during the 150 to 200 millisecond interval. A readiness potential appeared as expected just before participants moved their fingers, but the RP also appeared when participants vetoed their movement. According to Libet (1985), this indicates that the vetoing participants did intend to act, but withheld from moving just after they became aware of the decision to move. Libet concluded that it is possible to cancel an act produced by unconscious brain processes by using metaphysical freedom.

In my opinion, Libet's conclusion that the vetoing implies having metaphysical freedom does not directly follow from his results. He claims that the readiness potential that is present before the veto indicates that subjects intended to act. Subsequently, the actual veto would prove that metaphysical freedom could be used in order to withhold from moving. But I argue that the only conclusion that follows from Libet's results is that the readiness potential of participants planning to move and actually moving, is similar to the readiness potential of participants planning to move and withholding

from moving. I don't see any clue of metaphysical freedom in that. Furthermore, just as they determine actual movements, physical processes might also determine the veto beforehand (Wegner, 2003).

It turned out Libet was already aware of this problem and came up with a possible solution (Libet, 1999). He stated that the conscious veto is not a result of unconscious processes, but that it's a conscious *control process*. It is possible that the factors on which the veto is based develop unconsciously, but that the act of vetoing itself happens without any link to the preceding unconscious processes. Furthermore, Libet claims there is no experimental evidence against the possibility that the control process may appear without development by prior unconscious processes (p. 53). Thus, Libet postulates a new theoretical entity, namely the control process, and subsequently claims that it does not result from unconscious processes because there is no evidence against this claim. Here, Libet commits the reasoning fallacy of *appeal to ignorance*, which asserts that a statement is true, because it has not yet been proven false (Walton, 1996). By using this fallacy, one shifts the burden of proof to the opponents. The burden of proof, however, is on him, because he suddenly invents some sort of control process to explain why people can have metaphysical freedom. He should have supported this claim with empirical evidence.

To sum up, it is too soon to accept that people can have metaphysical freedom through the existence of a conscious veto that can override decisions that are made by unconscious brain processes. However, it is also too soon to draw the conclusion that metaphysical freedom does not exist. For one thing, there are a number of methodological problems that are leveled against Libet's experiments. In the next section, I review these criticisms. I argue that some are valid, and others can be questioned. The valid ones, however, indicate that Libet's investigations in themselves are inconclusive regarding the existence or in-existence of metaphysical freedom. The next paragraph will provide an overview of the most important shortcomings of Libet's paradigm. These criticisms focus on the relationship between the RP and the subjective awareness of the decision to act, as investigated by Libet.

Criticism on Libet's experiments

Recently, various scholars started to conduct studies that were designed to check whether Libet's findings could be replicated (Trevena & Miller, 2002; Trevena & Miller, 2009). In general they focused on the part in which Libet showed that unconscious brain processes (RP) related to a decision precede the awareness of that decision. They do not focus on the possibility of showing metaphysical freedom by using a veto.

Trevena and Miller (2002) used a methodology similar to Libet (1983). However, they asked participants to report the time of the *decision* to move, whereas Libet asked them to report the time of the *first awareness* of the decision to move. It turned out that in 20% of the trials, participants reported making a decision *before* the occurrence of a readiness potential. Trevena and Miller conclude that the readiness potential

does not precede the decision to move.

Libet (2002) wrote a review of the Trevena and Miller (2002) study in which he attacked their methodology as being the cause of their finding that participants reported making a decision (before) the occurrence of a RP. He argues that Trevena and Miller could not replicate his results because they asked participants to report the time of the *decision* to move and not the time of the *awareness* of the decision to move. Libet (2002) also points out that their conclusion is based upon the observation that participants reported making a decision before the occurrence of the readiness potential in 20% of the trials. Trevena and Miller did not elaborate on the 80% of the trials in which participants reported awareness *after* the RP. Furthermore, Libet argued that Keller and Heckhausen (1990), who used a methodology consistent with Libet's paradigm, did replicate his results.

Another confusing result of the Trevena and Miller (2002) study is that 40% of the participants reported the awareness of the decision after the action had been performed. Trevena and Miller explained this finding by stating that those participants might have simply been wrong about the time of their decisions. Yet, this then seems to indicate that the way in which Trevena and Miller measured the awareness of participants lacks precision. This lack of precision might also explain that in 20% of the trials participants reported awareness before the RP. This methodological imprecision seems to make the conclusions of Trevena and Miller unwarranted.

In another replication of Libet's study, Trevena and Miller (2009) asked 21 participants to perform the same task as in Libet's (1983) experiment. However, subjects were half of the time asked to always press a key, while during the other half of the time subjects were asked to sometimes press a key. In this way the researchers could monitor EEG activity at moments subjects did not make the decision to press a button. It turned out the EEG activity (RP) as reported by Libet was replicated in the present experiment. However, it appeared that this result was independent of whether subjects actually made the decision to press the button. The authors therefore conclude that the RP did not indicate the preparation of a movement, but that it possibly represents paying attention. These results imply that more accurate brain scanning methods than EEG are needed if we want to discover whether the brain activity is indeed related to a decision and not simply to paying attention.

Another criticism of Libet's experiments comes from Zhu (2003) who claims that Libet's paradigm contains two main flaws. First, Libet assumes that movements exclusively originate out of unconscious brain processes. However, subjects were asked beforehand to flex their fingers at some point in the future, which can be seen as facilitating a prior conscious intention to act before the experiment started. It can be argued that this critique does not pose any threat to Libet's results that unconscious brain processes precede the awareness of a decision. For the prior conscious intention to act could also be produced by unconscious brain processes that precede the conscious intention.

Zhu's second critique is that participants in Libet's study

did not report the awareness of the decision to move, but the awareness of an urge to move. An urge would be a feeling of wanting to move – and not a decision, because the urge has already been induced by the experimental instructions. Besides, if it turns out to actually be a decision and not an urge, then one could still argue that the decision of flexing a finger is extremely simple, because participants know in advance that they at some point will have to flex their fingers and they have to pick the hand with which they will do so in advance. So the only thing participants actually decide is the time of the flexing. If Libet wants to draw conclusions about the existence or in-existence of metaphysical freedom, then he will also have to design studies with more complex decision problems that involve conscious deliberation.

Finally, two other points of critique come from Klemm (2010). First, he indicates that the neurophysiological measures are not very adequate. Libet only measured EEG activity in the motor cortex and recent fMRI studies have indicated that decision-making is correlated to brain regions in the frontal cortex (Gold and Shadlen, 2007). Klemm's second critique is that the time of the awareness of a decision is measured through introspection. Participants are asked to indicate in retrospect where they thought the spotlight was located. This operationalization of the moment of awareness of a decision might not be very accurate.

To conclude, Libet's studies have indicated that unconscious brain processes related to a decision precede the awareness of that decision. Important limitations of Libet's studies are that the decisions might not even be a real decision and that some of his measures might be inaccurate. It is therefore too soon to draw any valid conclusions about the existence of metaphysical freedom. Future research should therefore try to come up with an experimental setup in which participants are required to make more complex decisions. Furthermore, more accurate neurophysiological measures should be used in order to detect whether other brain regions are also involved in the decision-making. Finally, new methods will have to be developed with which the time of the awareness of a decision can be determined more accurately. Most of the shortcomings of Libet's paradigm are accounted for by recently conducted studies. Therefore, the next chapter will provide an overview of those studies, together with other recent studies related to the relation between metaphysical freedom and phenomenological freedom.

Recently conducted studies

During the last decennium, a number of studies have been conducted within the Libet paradigm that tried to overcome the limitations of Libet's original studies (Lau et al., 2004; Soon et al., 2008; Fried, Mukamel & Kreiman, 2011). These studies tried to discover whether metaphysical freedom is illusory. Other recent studies did not look at the metaphysical kind of freedom but at phenomenological freedom.

The aim of the present chapter is to give an overview of the recently conducted studies about metaphysical freedom (within Libet's paradigm) and about phenomenological freedom. At the end of this chapter I will provide an answer

to the first aim of the present study: are there plausible reasons for the perspective that people do not have metaphysical freedom?

Metaphysical freedom

An important limitation of Libets experiments was that the neurophysiological measures were not very accurate. Libet used EEG measurements, which have high temporal resolution, but low spatial resolution (Huettel, Song & McCarthy, 2009). To overcome this limitation, Lau and his colleagues (2004) replicated Libets experiment but instead of EEG they used fMRI, which is a brain scanning technique with high spatial resolution. In line with Libet they found that the supplementary motor area is activated during the task, but they also found that activation of the pre-supplementary motor area (pre-SMA) precedes activation of the SMA. In fact, activation of the pre-SMA turned out to be related to the actual intention to move, while activation of the SMA was related to the preparation and execution of the movement. This study indicates that it is important to not only look at the SMA but also at the pre-SMA and perhaps even at other brain regions.

A limitation of the study by Lau and colleagues (2004) is that they did not measure whether the brain activation in the SMA and pre-SMA actually preceded the conscious awareness of the decision to move. Soon and his associates (2008) tried to overcome this limitation by also conducting an fMRI study within Libets paradigm, but this time including a better temporal measurement of the moment people became aware of their decision to move. The authors also planned to improve two other problems of Libets methodology. First, the time between the readiness potential (RP) and the reported awareness is only a few hundred milliseconds according to Libet. If the measurements would only be a little bit inaccurate, then this could lead to misplaced conclusions about the primacy of either brain activation or awareness of the decision. The second problem concerns the question of whether the brain activity is actually related to the subsequent decision, and does not simply relate to preparatory processes. In other words, can the brain activity be used to predict a choice between different types of behavior?

Soon and his colleagues (2008) solved these three problems by asking 36 participants to fixate at a screen on which every 500ms a different letter was shown. At some point, participants had to press either a left or a right button. They were completely free to decide when they wanted to do this. Participants were instructed to remember the letter that appeared on the screen at the moment they become consciously aware of choosing between the left and the right button. This method is a more accurate way of determining the actual moment of the awareness of a decision than the spotlight method used by Libet. Soon and colleagues claim it is easier to remember an actual letter than to retrospectively report the position of a moving spotlight. Brain activity was measured by using fMRI and the moment of the actual button press was recorded by using a computer. First, it turned out that the authors replicated Libets result that the SMA was activated around the moment of the decision to move. Furthermore,

they found that there was one brain area in particular that was predictive of whether the participants would press the left or the right button. Activity in the frontopolar cortex was present seven seconds before the subjects became aware of the decision. In some cases, the activity was present at even 10 seconds before the awareness. Soon and his colleagues (2008) were even able to use information about the frontopolar cortex activity pattern to predict the choice between the left and the right button with 60% accuracy (above chance level).

The study by Soon and his colleagues indicates that brain activity related to a decision is already present between 7 and 10 seconds before the actual awareness of the decision. This result implies that those simple decisions are actually the result of unconscious brain processes and not of our metaphysical freedom. However, an important limitation of this study is that the accuracy of the predictions is only 60%. In order to establish a real causal connection between brain activity and choices the decisions will have to be predicted with 100% accuracy. A possible explanation for the low accuracy of the predictions is the relatively low resolution of contemporary fMRI scans (Haynes, 2011).

Fried, Mukamel and Kreiman (2011) tried to overcome this limitation by using another scanning technique that looks at the single neuron level. The authors start by claiming that EEG recordings do not have high spatial resolution and fMRI studies do not have high temporal resolution. A single neuron recording, however, has both high spatial- and temporal resolution. Fried and his colleagues asked 12 epileptic patients to perform the classic task designed by Libet. They were able to look at individual neuron level because the epileptics had electrodes implanted in their brain because of their treatment. The electrodes were implanted in the Medial frontal cortex, a brain area associated with movement related decisions (the SMA also resides in this brain area). The study replicated the results of Libet by pointing out that cerebral brain activity in the SMA preceded the awareness of a decision to move with 1500 milliseconds. Furthermore, it turned out that activity in single neurons in the medial prefrontal cortex was predictive of the conscious intention to move on a single trial basis. The accuracy of these predictions was 80% at 700 milliseconds before participants became aware of their decision to move.

The authors conclude that our decisions are made by unconscious brain processes and that the awareness of those decisions is a product of the medial frontal cortex activity reaching a particular threshold. The theory that awareness of a decision only arises if a certain threshold of brain activity is reached has also been postulated by Gold and Shadlen (2007) and is in line with the results of a study by Matsushashi and Hallett (2008).

To sum up, it turns out that the methodological problems of the classic Libet experiments are largely solved by these recent experimental studies. More accurate brain scanning methods were used and it became even possible to predict the outcome of decisions with a prediction of 80%. This result, however, is still 20% below perfect accuracy, but future

improvements of brain scanning techniques might result into even more accurate predictions.

Furthermore, Soon and his colleagues (2008) improved the measure of the awareness of a decision by letting people remember a letter that appeared on a screen at the moment of the awareness. While this method certainly improves the spotlight method used by Libet, it is still based on introspection, which is not a very reliable measure of the actual moment of awareness (Khn & Brass, 2007). However, some of the studies described in this chapter showed that brain activity related to a decision is already present between 7 and 10 seconds before the awareness of that decision. It could be true that introspection is unreliable, but it would be very unlikely if the unreliability of introspection implies that participants made errors the size of multiple seconds.

The limitation of simple decisions has been improved a little by some of the studies described above. People now really had to choose between different actions and this is more complex than simply deciding when to flex your fingers. However, the decisions can still be considered as relatively simple and involve only limited reasoning. Therefore, future research should also focus on more complex decisions involving reasoning, such as moral dilemmas with two possible choices. It would be interesting to see whether it is possible to also predict the outcome of such decisions with high accuracy. One problem, however, with these kinds of complex decisions is that a moral dilemma usually involves reasoning and it will therefore be more difficult to ask participants at what time they became aware of making the decision. Determining the exact moment of becoming aware of making the simple decisions used in Libet's study is much easier because the decision suddenly arises in your mind, without a long process of conscious deliberation. Future research should try to invent a decision task that is complex, but still enables participants to easily identify the exact moment of making the decision.

It can be concluded that the studies described in this chapter indicate that simple decisions are produced by unconscious brain processes and the conscious awareness of these decisions arises later on, possibly as a result of brain activity in the medial frontal cortex reaching a threshold. These studies support the hypothesis that metaphysical freedom does not exist while making simple decisions. Future research should focus on increasing the prediction accuracy to 100% and discovering whether people also lack metaphysical freedom while making complex decisions.

Despite the conclusions above, we still have to explain the psychological fact that we all experience ourselves to be free while making decisions. As Libet (1999, p. 56) puts it: "A theory that simply interprets the phenomenon of free will as illusory and denies the validity of this phenomenal fact is less attractive than a theory that accepts or accommodates the phenomenal fact". Therefore, it is important to study phenomenological freedom and discover how it relates to metaphysical freedom. The next paragraph will focus on this question by describing some studies that tried to manipulate phenomenological freedom.

phenomenological freedom

Given that our experience indicates that we have free will (phenomenological freedom), and given that this free will is likely to be illusory on a metaphysical level, it is interesting to find out whether this phenomenological freedom is also simply a construction of our brains.

In order to find out whether that's true, Wegner and Wheatley (1999) asked 51 students to participate in an experiment in which they worked together with a confederate of which the participants thought it was another participant. The confederate and the participant were asked to watch a computer screen and both lay one hand on a computer mouse. They were requested to move the mouse in sweeping circles and once in every 30 seconds stop moving. After that, they had to indicate the extent to which they thought the stop was due to their own intentions. Participants had to wear headphones through which they heard music. The music started to play after the 30 seconds movement period and the participants were asked to stop the mouse from moving during the 10 seconds that the music was playing. Sometimes participants would also hear words through the headphones. Those words were used to prime thoughts about items that appeared on the computer screen. The confederate was required to sometimes stop the mouse movement on the item that was heard by the participant. The stop would be at different time intervals from when the word was heard: 30 seconds after, 5 seconds after, 1 second after, and 1 second before the participant heard the word. It turned out that, off all the forced stops, participants still thought they intended them in 52% of the cases. It turned out that participants perceived more intentionality as the spoken word was closer in time to the movement stop, but the experienced intentionality decreased as the word was spoken after the stop (Wegner & Wheatley, 1999).

The authors conclude that phenomenological freedom is an illusion created by the perception of a causal sequence between thought and action. People perceive themselves as the initiator of an action, while in fact someone else was the initiator. This illusion is enhanced by the presence of a thought that is related to the subsequent action. The conclusion that phenomenological freedom is actually a construction can also be connected to Libet's studies. In his experiments, people report the moment they become aware of a decision. At that moment, those people experience phenomenological freedom. It would be interesting to find out whether – besides the occurrence – also the timing of this phenomenological experience can be manipulated.

In order to find out whether that is possible, Isham and Banks (2009) have constructed an experiment similar to the classic experiment by Libet, but now a total of 8 subjects heard a beep as soon as they pressed the response button. The timing of this beep was manipulated in such a way that the beep was delayed with 5, 20, 40, or 60ms. The authors expected that a delayed beep would influence the judgment of the moment at which participants indicated that they became aware of their decision to press the button. It turned out that as the delay got longer, participants started to show

a delay in the time at which they became aware of deciding to press the button. In fact, the length of the delay explained 97% of the variance in the reported time of awareness. This implies that the exact moment at which people think that they become aware of making a decision is largely a reconstruction based on environmental cues. We do not perceive the moment we decide to act, but we infer it. These results are in line with Wegners (2003) ~~emph~~ priority principle, which states that the phenomenological freedom that people have at the moment of becoming aware of a decision is not based on actual awareness, but inferred from apparent cues. The priority principle also states that the extent to which people experience phenomenological freedom depends on the time between the intention and the intended behavior.

The two studies described above indicate that our phenomenological freedom can be seen as a construction. It can be manipulated in such a way that we experience that we are the authors of actions that we in fact did not even initiate. Furthermore, our experience of the time at which we have phenomenological freedom can also be altered, which also suggests that this kind of free will is a construction, most likely produced by our brains. Future research should try to discover the cerebral mechanisms that underlie the construction of our phenomenological freedom.

Discussion

The first aim of the present article was to find out whether metaphysical freedom exists. Libets classic studies indicated that brain activity related to a decision about a very simple movement was already present 550 milliseconds before people became aware of their decision to move. The methodology used by Libet had its limitations and most of them were solved by subsequent research. Thanks to better brain scanning devices, scholars were able to improve and extend Libets studies and it turned out that there is considerable support for the hypothesis that metaphysical freedom does not exist. Furthermore, some studies have indicated that phenomenological freedom is constructed. Future research could try to find out why we as human beings have developed the capacity to experience ourselves as freely acting agents, while in reality we do not have metaphysical freedom.

Given the numerous studies that support the hypothesis of the illusory nature of metaphysical freedom, scholars who still claim that metaphysical freedom exists, are pressed to come up with empirical evidence to support that claim. After all, they are postulating some sort of process through which people would be able to intervene the deterministic causal chain of nature. Introducing these kinds of processes – without empirical support – leads to a less parsimonious worldview and fall prey to Occams razor: without empirical support for that claim we would have to conclude that a theory without metaphysical freedom is more parsimonious and therefore should be accepted.

Despite the scientific research that supports the illusory nature of metaphysical freedom, there is still among laypeople a widespread belief in the existence of free will. In fact, a large survey taken in 36 countries indicated that 70% of the

respondents believe they have free will (International Social Survey Programme, 1998). It is perfectly understandable that so many people intuitively tend to think they have free will. After all, when making a choice we most of the time have a subjective experience of being able to choose different options.

However, I argue that this sense of freedom is not the metaphysical freedom that Libet and similar scholars are interested in. It can better be seen as phenomenological freedom and because it concerns freedom on an experiential level it does not tell us very much about our freedom in a deeper ontological sense. We have to be careful to not support ontological statements with phenomenological evidence. This distinction between ontology and phenomenology is very important and does not relate solely to the problem of free will. As I have argued in other work (Lodder, 2011) it is also linked to the question of whether the existence of the self is an illusion. In that debate people often use intuitions about the existence of a self as evidence for the ontological claim that there actually exists something like a self in the world.

Needless to say, the potential conflict between our intuitions and the real state of the world does not make our intuitions less interesting. People have those intuitions for a reason and it is the task of psychological research to find out what exactly these reasons are. If our intuitions about the world do not match reality then it is interesting to find out what process distorts our view of the world. Furthermore, the intuitions that people have about, for example, moral responsibility, are important from a judicial point of view. Jurisdiction in our society entails that a judge or a jury decides about the responsibility of someone that is suspected of having committed an unethical act. Those decisions are partly shaped by the intuitions that the judge or the members of the jury have about the extent to which people are morally responsible for their actions.

Given the conclusions of the first part of this thesis, it can be argued that our actions are completely determined by prior causes, described by the laws of nature and that a metaphysical kind of free will is an illusion. This conclusion might have influence on the way in which people think about moral responsibility, because can we still be held morally responsible for our actions, if it turns out that those actions are completely determined by factors other than our (illusory) metaphysical freedom? Therefore, the aim of the next chapter is to investigate whether moral responsibility is still possible if metaphysical freedom is illusory.

Moral Responsibility

For a long time, philosophers dominated debates about the compatibilism of determinism with free will and moral responsibility. In order to support their statements, some philosophers would refer to claims about the intuitions people have about compatibilism. Robert Kane and Peter Strawson, two well-known philosophers in this debate, have claimed that people intuitively think that determinism and moral responsibility are incompatible (Kane, 1999, p. 218; Strawson, 1986, p. 30). This would imply that if determin-

ism is true, most people would think that people who commit a crime are no longer responsible for their actions.

The claims made by Kane and Strawson, however, lack any empirical support (Nichols & Knobe, 2007). Yet, recently, a new field of research is emerging within philosophy, coined experimental philosophy. Scholars working in this field are conducting scientific experiments in order to support their philosophical claims. This main goal of this chapter is to give an overview of the studies conducted within the field of experimental philosophy that focus on the compatibilism of determinism with free will and moral responsibility.

Intuitions about compatibilism

A number of experimental philosophers started to study the intuitions people have about the compatibilism of determinism with free will and moral responsibility. Nahmias and his colleagues (2005) were the first to come up with an experiment that had to provide an answer to this question. They asked 21 participants to read a scenario about a world in which a supercomputer is able to predict everything that will happen in the future. One day, the computer predicts that some person will rob a bank at a particular time in the future and the prediction turns out to be true. Subsequently the participants are asked to indicate whether they think that the person robbed the bank out of his free will and whether they think this person is blameworthy for his acts. It turned out that 76% of the participants indicated that the person robbed the bank out of his own free will. Furthermore, 83% of the participants judged the person to be blameworthy of his actions.

The authors conclude that most people do not judge determinism to be incompatible with free will and moral responsibility. Compatibilistic judgments were made three times as much as incompatibilistic judgments. There is, however, an important limitation of the study described above. Most participants were laypeople philosophically and were therefore not familiar with philosophical concepts. It might be possible that participants did not interpret the scenario as being deterministic, but perhaps as being fatalistic. Therefore, Nahmias and his colleagues (2006) wanted to replicate the previous experiment but now with two different operationalizations of a deterministic scenario.

The first scenario was one in which people live in a universe that is recreated over and over again, in such a way that every time the universe is recreated the same things will happen. The second scenario is one in which all the actions of people are determined completely by either genes or the environment. Subjects are again asked to indicate whether a person that committed a crime (stole something) did this out of his own free will and whether this person can be held morally responsible for his deeds. It turned out that 66% (scenario 1) and 76% (scenario 2) of the participants indicated that the person acted out of his own free will. Furthermore, 77% (scenario 1) and 60% (scenario 2) judged that this person was also blameworthy for his deeds. There were no significant differences between judgments in the two scenarios. Participants who thought that the person could still have

free will, also tended to think that the person was morally responsible for his actions.

Again, these results indicate that people show Compatibilistic intuitions concerning the possibility of free will and moral responsibility in a deterministic world. These results are in contrast with the statements of Kane and Strawson mentioned in the introduction, who said that people have incompatibilistic intuitions. These findings therefore underline the importance of conducting experimental research within philosophy.

The results of the two studies by Nahmias and his colleagues (2005; 2006) show that *not everyone* has compatibilist intuitions. There still is some variation present because on average about 75% of the people gave compatibilist answers. Therefore, some scholars started to wonder if there are other factors influencing whether people show compatibilist intuitions or not. One possible factor could be the emotions that are triggered by reading scenarios about immoral acts (Lerner, Goldberg & Tetlock, 1998; Smart en Loewenstein, 2005).

In order to discover whether affective processes influence the intuitions people have about compatibilism, Nichols and Knobe (2007) presented 41 participants with two scenarios. A concrete scenario was designed in order to induce an affective response and a more abstract scenario was designed to induce a less affective response. Their prediction was that participants in the concrete condition would show more compatibilist responses than people in the abstract condition, who in turn are expected to show more incompatibilist responses. The scenario in the concrete condition was about a man who wanted to be with his secretary and therefore killed his whole family by burning down his house. Participants were given a description of a deterministic universe and were asked to judge whether the man can be held morally responsible for his actions in such a universe. In the abstract condition participants were just given the description of a deterministic universe and were only asked whether people in such a universe could be held morally responsible for their actions. It turned out that in the concrete scenario, 72% of the participants gave *compatibilist* responses, while in the abstract scenario, 86% of the participants gave *incompatibilist* responses.

The authors conclude that people intuitively can have either compatibilist or incompatibilist responses and that the responses they tend to show depend on affective psychological processes. The emotions that were triggered in concrete moral scenarios are thought to reduce the extent to which people are able to apply their incompatibilist intuitions. This result was replicated by Nahmias, Coates, and Kvaran (2007), but they also found that people not only show more compatibilist intuitions if the scenarios are described in concrete terms, but also if the deterministic universe is described in more concrete terms.

The researchers provided their participants with two different scenarios. In the *concrete deterministic* scenario people are told that decision making is determined completely by certain psychological intentions or desires, while in the

abstract deterministic scenario, people are told that neurochemical processes and genes completely determine our decisions. It turned out that in the concrete condition 88,6% of the participants showed compatibilist responses, compared to only 40,7% in the abstract condition.

These results imply that it is not only the concreteness of the described scenario, but also the concreteness with which determinism is described, that determine the extent to which people show compatibilist intuitions. This could imply that stories that are more familiar to us lead us to judge in a more compatibilist way. Therefore, Roskies and Nichols (2009) investigated whether the closeness of the world in which the scenario takes place influences whether the response is compatibilist or incompatibilist.

In order to answer that question, 76 participants were randomly divided over two conditions. In the *actual* condition participants were presented with a description about determinism and in the *alternate* condition this description was framed as if it concerned the determinism of beings in another universe. Participants in both conditions were asked to indicate – based on the presented information – the extent to which they thought a person was morally responsible for his actions. It turned out that participants in the actual condition gave significantly higher ratings of freedom and moral responsibility than people in the alternate condition.

These results imply that stories about our own world lead to more compatibilist responses than stories about another universe. The interpretation of Nichols and Knobe (2007) that concrete scenarios involve more affective processes could also be applied to the findings of Roskies and Nichols (2009). It could be that stories about our own universe induce affect to a greater extent than stories about other universes. Given that affective psychological processes play a role in the intuitions people have concerning compatibilism, it is interesting to find out whether these intuitions are influenced by other psychological processes, such as personality traits. It has been shown that people with high extraversion are more socially minded than people with low extraversion (Ashton, Lee, and Paunonen, 2002). Because of this social focus of extraverts, one could expect that they will have a higher chance than introverts of assigning someone responsible for harming people in a deterministic world.

In order to test this, Feltz and Cokely (2009) presented 58 participants with a scenario about a man who kills his wife. Participants were asked to indicate to what extent they judged that the man was acting out of free will and was morally responsible. Participants were also asked to fill out a mini version of the Big V personality test. The authors were only interested in the extraversion dimension of the Big V. It turned out that people with high extraversion scores, tended to judge the man as being more responsible and having more free will, compared to people with low extraversion scores.

These results imply that personality differences on the extraversion trait influence the intuitive judgments about the compatibility of determinism and free will or moral responsibility. The studies described in this paragraph indicate that there is no single set of folk intuitions, but that the kind of

intuitions one has depends on the personality and the concreteness with which a story and determinism are described. People tend to have compatibilist intuitions if they are more extraverted and if the story and determinism are described in concrete terms, while people have more incompatibilist intuitions if they are introvert or the story and determinism is described in more abstract terms.

According to Roskies and Nichols (2009) these results imply that if scientists would someday denounce that people do not have metaphysical freedom and that all our actions are the results of the prior state of the universe combined with the laws of nature, that this would not change very much about the way we as human beings live our lives. We still would mostly think that people could have moral responsibility for their actions, even if those people do not have metaphysical freedom. This statement is based upon conclusions about what people say they would do and not about what people actually do. The next paragraph will describe some studies in which scholars looked at the changes in behavior of people who were told that metaphysical freedom did not exist.

Determinism and Behavior

Some philosophers have argued that personal responsibility will decrease if people think they have no metaphysical freedom in their actions (Smilansky, 2000). Vohs and Schooler (2008) investigated whether people would behave differently if scientist finally declared that mankind does not have metaphysical freedom. They divided 30 participants over two conditions. In the *anti-free will* condition they were asked to read a text about scientists who declared metaphysical freedom to be an illusion. In the *control* condition participants were asked to read a text about consciousness not related to the problem of free will. Subsequently, participants were asked to solve some math problems on a computer. They were told that the computer had a glitch and that the answer on the question would at some point appear on the screen, but not if the participants would press the spacebar as soon as the question appeared. Participants were told that the experimenter would not know whether they looked at the answers, but that they should try to answer the questions honestly. It turned out that participants in the anti-free will condition cheated significantly more frequent than participants in the control condition. The authors conclude that priming a belief in determinism encourages cheating behavior. It might be interesting to look whether people do not only show an increase in immoral behavior, but also a decrease in moral behavior.

Baumeister, Masicampo and DeWalt (2008) tried to answer this question by looking at the influence of the non-existence of metaphysical freedom lead to a decrease in helpfulness. They randomly divided 70 participants over three conditions. In the *determinism* condition, participants were asked to read statements relating to determinism, such as Like everything else in the universe, all human actions follow from prior events and ultimately can be understood in terms of the movement of molecules. In the *free will* conditions, participants were asked to read statements about metaphys-

ical freedom (I am able to override the genetic and environmental factors that sometimes influence my behavior) and in the *neutral control* condition participants were presented statements about scientific facts (Oceans cover 71% of the earth's surface). Subsequently, participants were presented with six scenarios about the possibility of helping someone (such as giving money to a homeless person) and were asked to indicate the likelihood with which they would help the person in each scenario. It turned out that participants in the determinist condition showed a significantly lower likelihood of helping than participants in the free will and neutral control conditions. The latter two conditions did not show a significant difference.

The authors conclude that prosocial behavior is reduced as a result of a belief in determinism. They explain this finding by postulating that people who were primed with a belief in determinism have a lower motivation to exert self-control and therefore tend to respond more automatically. An important limitation of this study (Baumeister, Masicampo & DeWall, 2008) and the one conducted by Vohs and Schooler (2008) is that the results do not indicate that people will change their behavior in the long run. The studies only point out that people behave more unethically only a few minutes after reading passages about the illusory nature of free will. A more objective way of studying the influence of a disbelief in free will on behavior would be to go out in the field and observe the actual behavior of people that have acknowledged their disbelief in metaphysical freedom.

Another limitation of the studies described above is that the participants might have confused determinism with fatalism. Fatalism means all events in the universe are destined to happen, independent of the causal sequence that preceded the event. In a fatalistic universe, everything that is meant to happen will eventually happen, independent of motivations, desires or intentions. If people would mistakenly interpret determinism in this way, then this could explain the increased cheating and decreased helping, because people might think that everything they would do was already predestined. Future research should therefore design experiments in which determinism is defined while preventing that people think that the future is predetermined.

The second aim of this thesis was to find out whether people can still be morally responsible if they do not have metaphysical freedom. The experiments described in this chapter indicated that the intuitions people have about this question do not point into one direction. First, it turned out people have compatibilist intuitions if they have extravert – rather than introvert – personalities and if the definition of determinism and the stories they have to judge are described in concrete, rather than abstract terms. These results, however, are not completely in line with the results of the behavioral experiments. It turns out that if people are primed with a belief in determinism they start to cheat more and show less helpful behavior. This implies that people tend to think that they lose moral responsibility if determinism is true. This supports the conclusion that people intuitively think that moral responsibility is *incompatible* with determinism.

These results are however not completely valid, because the described studies have some methodological problems.

The experiments described in this chapter, shed light on the intuitions and beliefs people have about the compatibility of determinism and moral responsibility. It is not possible, however, to use this data in order to inquire in whether people actually *are* morally responsible, given the illusory nature of metaphysical freedom. Therefore, in the general discussion I will try to answer this question by using philosophical reasoning.

General discussion

It turns out that there are reasons to assume that metaphysical freedom is illusory and, hence, phenomenological freedom is a construction. The studies described in the previous chapter concern phenomenological data and cannot be used in order to draw any conclusions about the ontological question of whether people can still have moral responsibility, given the illusory nature of metaphysical freedom. Therefore, I will now try to answer this question based on philosophical reasoning.

Within the compatibilism debate, some philosophers equate metaphysical freedom with moral responsibility in such a way that you need the former to have the latter (Zhu, 2003; Strawson, 1986; Mele, 2006). Furthermore, the studies about the intuitions of laypeople about compatibilism indicate that people who think that determinism is compatible with metaphysical freedom also tend to think that determinism is compatible with moral responsibility. This view might also be the result of the assumption that metaphysical freedom is necessary in order to have moral responsibility. These views automatically results into only one type of compatibilism, namely the compatibilism of determinism with both moral responsibility *and* metaphysical freedom. In my view, however, there are two different types of compatibilism:

- (1) compatibilism of determinism and metaphysical freedom
- (2) compatibilism of determinism and moral responsibility

Based on the conclusions in the first part of this thesis, I would like to argue that the first type of compatibilism does not hold: determinism and metaphysical freedom are incompatible. Some philosophers, however, still argue despite the scientific evidence - that metaphysical freedom is compatible with determinism. First, I will summarize the main argument that are advanced in support for this position and indicate why I think this argument does not hold. Subsequently, I will find out whether there are reasons to assume that the second type of compatibilism is plausible.

Daniël Dennett (2003) is a contemporary philosopher who thinks that there still is room for metaphysical freedom, despite the conclusions of the experiments described in the first part of this thesis. Dennett states that it is true that unconscious processes in our brains determine our actions, but at the same time those actions are still *our* actions. If our brains make a decision, then this decision is still *our* decision and therefore we have free will. According to Dennett, free will is also something that evolves with us. As human beings

evolved, they developed the capacity to use language and reason while guiding their behavior. This increase in the extent to which we can perform complex behavior, makes us freer. As the title of Dennett's book indicates: freedom evolves (Dennett, 2003).

My main criticism on this kind of argument is that it does support the compatibilism of determinism with moral responsibility (as I will argue below), but it does *not* support the compatibilism of determinism with metaphysical freedom. It does not support this type of compatibilism because Dennett changes the definition of free will in such a way that it becomes compatible with determinism. In his definition of free will, Dennett includes unconscious processes that are deterministic in nature. According to Dennett, unconscious processes that guide our behavior are just as much part of us as our conscious thoughts. Furthermore, our freedom evolves together with the degree in which we can show complex behavior as human beings. This kind of freedom, however, can be seen as *behavioral freedom* and not as metaphysical freedom. If freedom evolves then this implies that free will increases if human beings become capable of more complex and diverse forms of behavior. This concerns the question of what types of behavior are possible, which relates to the definition of *behavioral freedom*. Metaphysical freedom concerns the question of whether human beings can perform actions that are not the result of a preceding deterministic process. Dennett does not provide an answer to this question because he talks about behavioral freedom instead of metaphysical freedom and therefore his work does not support the thesis that determinism is compatible with metaphysical freedom. It is therefore still plausible to assume that the first type of compatibilism – between determinism and metaphysical freedom – does not hold.

In order to discover whether the second type of compatibilism between determinism and moral responsibility holds, it is important to first make a distinction between *being* morally responsible in an ontological sense and holding someone morally responsible in a conventional sense. The former concerns the question of whether people actually *are* morally responsible, given the illusory nature of metaphysical freedom. The latter concerns the question of whether it is still possible to hold people responsible even if they are not responsible in an ontological sense. I am going to argue that given the illusory nature of metaphysical freedom, people are not morally responsible in an ontological sense, but they can be *held* morally responsible for their actions because they are part of a society with certain conventions.

It is first important to make a distinction between two kinds of legal systems (Green and Cohen, 2004): A retributive legal system and a consequentialist legal system. In a *retributive* system people are mainly punished because they have to get what they deserve based on their actions. A consequentialist system uses punishment only as an instrument in order to achieve social welfare in the future. The most salient difference between the two legal systems is that a retributive system focuses on the immorality of a *person* who has to be punished, while a consequentialist system focuses

on the immorality of an *act*, that has to be corrected.

As mentioned earlier, some philosophers equate moral responsibility with metaphysical freedom in such a way that you need the freedom to have the responsibility. In my opinion, this kind of reasoning is entangled with a retributive legal system, because the focus is on the people (that have to be free in order to be responsible), instead of on the acts. On the other hand, if one takes a consequentialist perspective, then it doesn't matter if a person has metaphysical freedom or not. The only thing that matters is that the person has committed a crime for which he or she has to be corrected.

Smilansky (2011) is one of the philosophers that equates moral responsibility with free will and concludes as a true incompatibilist that: if determinism is true, then both free will and moral responsibility are not possible. Smilansky argues that the position of hard determinism, which implies that people do not have free will and are therefore not morally responsible for their actions, requires us to abandon punishing immorally behaving individuals because they cannot be seen as morally responsible. If hard determinism is true, then people do not control their actions and do therefore not deserve to suffer for them.

Strawson (1986) has a view similar to that of Smilansky (2011). He claims that, according to the basic argument, people cannot be held fully morally responsible for their actions, because actions arise out of mental states and people are not fully responsible for the mental states they have and therefore for the way they behave. Every act and every thought in our mind is the result of features for which we are ultimately not responsible, such as genes and our environment. It is therefore exactly as just to punish someone for an action, as it is to punish someone for his or her hair color (Strawson, 1986).

First of all, I do not agree with Strawson's argument, because having a particular hair color is not an act for which we can be held morally responsible within a particular society, whereas killing someone is. Therefore, the person who killed someone should be responsible for this act and has to be corrected in order to never commit that crime again. Strawson and Smilansky are both right that people themselves are largely not the cause of who they are and are therefore not morally responsible in an ontological sense; our genes and the environment shape who we become to a very large extent. But what counts is that a person is part of a society that views certain acts as immoral. Therefore, from a conventional point of view, the only thing that is relevant is that this person acted in a morally wrong way and must therefore be corrected. Given the illusory nature of metaphysical freedom, it is not true that the people are morally responsible, but it could be argued that they can still be held morally responsible because if they committed an immoral act. In this case the focus is on the act instead of on the person, which is consistent with a consequentialist legal system.

In my opinion, the replacement of a retributive legal system with a consequentialist legal system, will lead to a more objective legislation. Because if we have no metaphysical freedom and every act is therefore the result of forces we do not control, then retributive punishment does no longer make

sense. Just like it doesn't make sense to punish a machine for being defective. The machine and the person only have to be corrected in such a way that they become determined to no longer perform this defective behavior (I am not reducing humans to machines here, because humans are at the moment capable of much more complex behavior than machines, but in essence the underlying principles that govern their behavior are both deterministic in nature).

Thus, moral responsibility as a convention can exist without having metaphysical freedom. People are responsible for their actions to the extent that they are the human beings that are part of a society and can be shaped in a way that lets them commit immoral acts and this will have to be corrected in order to achieve a more peaceful society. My views on this matter can be explained more clearly by considering the following three hypothetical scenarios:

- (1) A man kills another man because of a brain tumor
- (2) A man kills another man because of a psychosis developed due to being sexually abused as a child
- (3) A man kills another man because he enjoys doing those kind of things

Most people would probably think that the man in the first scenario is least responsible for his actions. But I propose a view in which this man is fully responsible to the extent that he is a human being who acted in a way that is not morally permitted. Therefore, he has to be corrected in order that he will never act in that way again. The way in which he has to be corrected depends on the identified causes of his actions. So in this case the tumor should be removed and after that it has to be determined whether the man is normal again so he can return in society.

Most people will probably think that the man in the second scenario is not fully responsible for his actions because of his psychosis. Again, in my view this man is fully responsible to the extent that he as a human being killed another man. Therefore, he has to be corrected so that he will never do it again. The way in which he has to be corrected depends on the reasons he had for his actions. So in this case he should be treated for his psychosis. If the treatment succeeds and the man gets cured of his psychosis (without having to use any medication in the future) then the determinants of the murder have been corrected and the man can return in society. If treatment does not succeed, then this man will have to be watched over until treatment succeeds or until he dies.

Most people will probably think that the man in the third scenario should be sent to hell or even worse, just because there seems to be no psychological or medical 'label' we can give this man. I agree that this person can be held morally responsible for his actions, but he is in the same way responsible as the previously mentioned examples. He as a human being has performed an action that is morally not permitted according to our conventions and he has to be corrected for that. The way in which he has to be corrected depends on the reason he had for his actions. In this case he seems to enjoy the killing, so this indicates that there is something deeply rooted in him that causes him to have absolutely no empathy

with his victims. Perhaps it is possible that he has some kind of disorder or abnormality in his brain, of which we are not aware that it exists. In that case we should lock the man up until we identify his problem and are able to treat him.

In my opinion, all three persons are equally responsible for their actions. However, they acted out of different reasons and therefore the treatment or correction of the person who enjoys the killing will most likely cost more effort than the treatment of the person with the tumor. But in the end, both will have to be corrected in such a way that they do not kill people any longer. One could, however, wonder to what extent it is possible to determine whether a person is completely cured or treated and does not have the risk of killing anyone again in the future. But this is a different question and goes beyond the scope of the present article.

All three cases illustrate that those men acted in a certain way because of their genetic/biological/neurological constitution (nature) combined with their past experiences and environmental influences (nurture). In each case the existence of metaphysical freedom is not necessary for letting those men be held morally responsible for their actions according to the conventions of our society. Furthermore, a brain tumor turns out to be only a special case of a physical event that determines our thoughts and actions. It is in essence similar to the neurological characteristics of the person who committed the murder because he likes it.

Note that another scenario in which a man was *forced* to kill another man would in my view imply that the forced man is *not* responsible for his actions. The difference between this- and the other three scenarios is that being forced is an exogenous determinant, while the other three scenarios involved an endogenous determinant. Endogenous determinants arise within a person and therefore the correction will have to be applied to that person. Exogenous determinants, however, do not arise within a person, but serve as an external limit to the *behavioral freedom* of a person. Correcting a person for an exogenous determinant (such as being forced to kill) does not make sense, because what has to be corrected for in the person if the determinant of the behavior was external?

I believe that dismissing the notion of metaphysical freedom leads to a better justice system that is purely based upon correcting the causes of an immoral act and thereby changing the determinants of immoral behavior in such a way that the person will start acting morally. Given the illusory nature of metaphysical freedom, it is no longer necessary to apply punishment if the immoral behavior of a person is already corrected. For why would we punish a person any longer if we know that he or she is completely treated and will therefore never perform immoral acts again?

To conclude, I think there are good reasons to assume that metaphysical freedom is illusory. Despite this conclusion, I still think it is possible to keep people morally responsible for their actions in a conventional way, because they are human beings that are part of a society. If their actions turn out to be evaluated as immoral according to the norms of the society which those individuals are part of, then this immoral

behavior will have to be corrected, but which correction is most suitable depends on the actual determinants of the immoral behavior. I sincerely hope that this thesis will aid a shift towards a more objective, honest and consequentialist legal system and consequently a more peaceful society.

I would like to end this thesis with a quote of Albert Einstein (Planck, 1932). This quote perfectly reflects my own view on the problem of free will:

Honestly, I cannot understand what people mean when they talk about the freedom of the human will. I have a feeling, for instance, that I will something or other; but what relation this has with freedom I cannot understand at all. I feel that I will to light my pipe and I do it; but how can I connect this up with the idea of freedom? What is behind the act of willing to light the pipe? Another act of willing? Schopenhauer's words 'Man can do what he will, but he cannot will what he wills,' accompany me in all situations throughout my life and reconcile me with the actions of others, even if they are rather painful to me. This awareness of the lack of free will keeps me from taking myself and my fellow men too seriously as acting and deciding individuals, and from losing my temper.

References

- Banks, W. P., and Isham, E. A. (2009). We infer rather than perceive the moment we decided to act. *Psychological Science*, 20, 17-21.
- Baumeister, R., Masicampo, E. J., and DeWall, N. (2008). Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin*, 35, 260-268.
- Dennett, D. C. (2003). *Freedom Evolves*. New York: Penguin.
- Feltz, A. and Cokely, E. T. (2009). Do judgments about freedom and responsibility depend on who you are? Personality differences in intuitions about compatibilism and incompatibilism. *Consciousness and Cognition*, 18, 342-350.
- Feltz, A., Cokely, E. T. and Nadelhoffer, T. (2009). Natural Compatibilism versus Natural Incompatibilism: Back to the Drawing Board. *Mind and Language*, 24, 1-23.
- Fried, I. Mukamel, R., & Kreiman, G. (2011). Internally Generated Preactivation of Single Neurons in Human Medial Frontal Cortex Predicts Volition. *Neuron*, 69, 584-562.
- Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 23, 829-839.
- Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annual Review of Neuroscience*, 30, 535-574.
- Huettel, S. A., Song, A. W., McCarthy, G. (2009). *Functional Magnetic Resonance Imaging* (2 ed.), Massachusetts: Sinauer, p. 449.
- International Social Survey Programme. (1998). Religion II. Retrieved December 26, 2006, from Gesis.
- Inwagen, P. van (1983). *An Essay on Free Will*. Oxford: Oxford University Press.
- Kane, R. (1999). Responsibility, luck and chance: Reflections on free will and indeterminism. *Journal of Philosophy*, 96.
- Keller, L. and Heckhause, H. (1990). Readiness potentials preceding spontaneous motor acts: Voluntary vs. involuntary control. *Electroencephalography and Clinical Neurophysiology*, 76, 351-361.
- Klemm, W. R. (2010). Free will debates: Simple experiments are not so simple. *Advances in Cognitive Psychology*, 6, 47-65.
- Kornhuber, H. H. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv Physiologie*, 284, 17.
- Khn, S., and Brass, M. (2007). Retrospective construction of the judgment of free choice. *Consciousness and Cognition*, 18, 12-21.
- Lerner, J. S., Goldberg, J. H. and Tetlock, P. E. (1998). Sober second thought: the effects of accountability, anger, and authoritarianism on attributions of responsibility. *Personality and Social Psychology Bulletin*, 24, 563
- Libet, B., Wright, E. W., and Gleason, C. A. (1982). Readiness-potentials preceding unrestricted 'spontaneous' vs pre-planned voluntary acts. *Electroencephalography and Clinical Neurophysiology*, 54, 322-335.
- Libet, B., Gleason, C. A., Wright, E. W., and Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness potential). *Brain*, 106, 623-642.
- Libet, B. (1985). Unconscious cerebral initiative and the role of conscious will in voluntary action. *The behavioral and brain sciences*, 8, 529-566.
- Libet, B. (1999). Do we have free will? *Journal of Consciousness Studies*, 6, 47-57.
- Libet, B. (2002). The timing of mental events: Libets Experimental Findings and their implications *Consciousness and Cognition*, 11, 291-299.
- Lau, H. C., Rogers, R. D., Haggard, P., and Passingham, R. E. (2004). Attention to Intention. *Science*, 303, 1208.
- Lodder, P. (2011). VRT-2: Is het zelf een illusie? Retrieved May 25, 2012, from [http://psychologyincludes.edu.fmg.uva.nl/owi/psychological_methods/Lodder\(2011\).pdf](http://psychologyincludes.edu.fmg.uva.nl/owi/psychological_methods/Lodder(2011).pdf)
- Mele, A. R. (2006, p. 3). *Free will and luck*. Oxford: Oxford University Press.
- Miller, J. S. and Feltz, A. (2009). Frankfurt and the folk: An experimental investigation of Frankfurt-style cases. *Consciousness and Cognition*, 20, 401-414.
- Nahmias, E. A., Morris, S., Nadelhoffer, T., and Turner, J. (2005). Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility, *Philosophical Psychology*, 5, 561-584.
- Nahmias, E. A., Morris, S., Nadelhoffer, T., and Turner, J. (2006). Is Incompatibilism Intuitive?, *Philosophy and Phenomenological Research*, 76, 321-346.

nomenological Research, 1, 28-53.

Nahmias, E. A., Coates, D. J., and Kvaran, T. (2007). Free Will, Moral Responsibility and Mechanism: Experiments on Folk Intuitions, *Midwest Studies in Philosophy*, 31, 214-242.

Nichols, S., and Knobe, J. (2007). Moral Responsibility and Determinism: The Cognitive Science of Folk Intuitions, *Nous*, 4, 663-685.

Planck, M. (1932). *Where is science going?* New York: W. W. Norton & Company, 201.

Roskies, A. L., and Nichols, S. (2009). Bringing moral responsibility down to earth. *Journal of Philosophy*, 8, 371-388.

Singer, I. B. (1997). *Isaac Singer's Promised City*. Retrieved May 25, 2012, from City-Journal.

Sarkissian, H., Chatterjee, A., de Brigard, F., Knobe, J., Nichols, S., and Sirker, S. (2010). Is belief in free will a cultural universal? *Mind and Language*, 25, 346-358.

Searle, J. R. (1983). *Intentionality: an essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Smart, D., and Loewenstein, G. (2005). The devil you know: the effects of identifiability on punitiveness. *Journal of Behavioral Decision Making*, 18, 311-318.

Smilansky, S. (2000). *Free Will and Illusion*. Oxford: Oxford University Press.

Soon, C. S., Brass, M., Heinze, H., & Haynes, J. (2008). Unconscious determinants of free decisions in the human brain. *Nature Neuroscience*, 11, 543-545.

Strawson, G. (1986). *Freedom and Belief*. Oxford: Oxford University Press.

Trevena, J. and Miller, J. (2009). Cortical Movement Preparation before and after a Conscious Decision to Move. *Consciousness and Cognition*, 11, 162-190.

Trevena, J. and Miller, J. (2009). Brain preparation before a voluntary action: Evidence against unconscious movement initiation. *Consciousness and Cognition*, 19, 447-456.

Vohs, K. D., and Schooler, J. W. (2008). The value of believing in free will: Encouraging a belief in determinism increases cheating. *Psychological Science*, 19, 49-54.

Walton, D. G. (1996). *Arguments from Ignorance*. Pennsylvania: Pennsylvania State University Press.

Wegner, D. M. (2003) *The Illusion of Conscious Will*, Cambridge: MIT Press.

Wegner, D. M., & Wheatley, T. (1999). Apparent Mental Causation: Sources of the Experience of Will. *American Psychologist*, 7, 480-492.

Wegner, D. M., Sparrow, B., & Winerman, L. (2004). Vicarious Agency: Experiencing Control over the Movements of Others, 6, 838-848.

Woolfolk, R. L., Doris, J. M., and Darley, J. M. (2006). Identification, situational constraint and social cognition: studies in the attribution of moral responsibility. *Cognition*, 100, 283-301.

Zhu, J. (2003). Reclaiming Volition. *Journal of Consciousness Studies*, 10, 61-77.