

# Control and Practical Moral Responsibility

---

*-A hard incompatibilist solution to the free will problem that doesn't require a revolution of our reactive attitudes*

## **Kontroll og Praktisk Moralsk Ansvar**

*-En løsning på fri vilje problemet fra hard inkompatibilisme, som ikke krever en revolusjon av våre reaktive holdninger*



**Navn: Maria Gamman**

**Veileder: Espen Gamlund og Mette Kristine Hansen**

**FILO350  
Mastergradsoppgave i filosofi  
Institutt for filosofi og førstesemesterstudier  
Universitetet i Bergen  
Høst 2013**

Thank you very much to Espen Gamlund and Mette Kristine Hansen for outstanding supervision and a lot of help.

Also, thanks to everyone who has bothered and been bothered to debate free will with me- Jøran Gamman, Hilde Gamman, Ingeborg Storesund, Linn Skagestad, my fellow students at UiB, ++.

## Contents

<b>Kontroll og Praktisk Moralsk Ansvar</b>	<b>2</b>
<b>0.0. Abstrakt på norsk</b>	<b>6</b>
<b>0.1. Abstract in English</b>	<b>7</b>
<b>0.2. Pretext</b>	<b>8</b>
<b>0.3. Introduction</b>	<b>9</b>
<b>1. Part one: the free will debate and favoring the position of hard incompatibilism</b>	<b>12</b>
<i>1.1. Introduction to part one</i>	<i>12</i>
<b>1.2. The free will debate</b>	<b>13</b>
<b>1.3. Freedom</b>	<b>14</b>
<i>1.3.1. Different philosophical understandings</i>	<i>14</i>
1.3.1.1. The garden of forking paths.	14
1.3.1.2. The source model of control.	14
1.3.1.3 Kane's five freedoms	15
<i>1.3.2. Common-sense</i>	<i>16</i>
<i>1.3.3. Intuition and experience</i>	<i>18</i>
<i>1.3.4 Control</i>	<i>20</i>
<b>1.4. The relationship between freedom and moral responsibility</b>	<b>22</b>
<b>1.5. The idea of determinism</b>	<b>23</b>
1.5.1. Confusions about determinism	23
1.5.2. Most philosophers endorse determinism	24
<b>1.6. Libertarianism in general</b>	<b>25</b>
<i>1.6.1. Some problems with general libertarianism</i>	<i>27</i>
<i>1.6.2. The best account of libertarianism</i>	<i>28</i>
<b>1.7. Consciousness, subjectivity and the free will debate</b>	<b>30</b>
<i>1.7.1. Consciousness</i>	<i>30</i>
<i>1.7.2. Subjectivity</i>	<i>31</i>
<b>1.8. Compatibilism</b>	<b>33</b>
<i>1.8.1. Classical compatibilism</i>	<i>33</i>
1.8.1.1. Van Inwagen and the consequence argument	34
<i>1.8.2. Modern compatibilism</i>	<i>36</i>
1.8.2.1. Modern compatibilists strategies for refuting the consequence argument	37
<i>1.8.3. Compatibilism according to Daniel Dennett</i>	<i>38</i>
1.8.3.1. Dennett's Evitability	40
<i>1.8.4. PAP and Frankfurt</i>	<i>41</i>
<i>1.8.5. Manipulation arguments</i>	<i>42</i>
1.8.5.1. Compatibilist replies to manipulation arguments	43
<b>1.9. Hard determinism</b>	<b>44</b>
<b>1.10. Hard incompatibilism</b>	<b>46</b>
<i>1.10.1. The hard incompatibilism of Pereboom</i>	<i>47</i>
<b>1.11. Conclusion of part one</b>	<b>48</b>
<b>2. Part Two: moral implications of hard incompatibilism</b>	<b>49</b>
<b>2.1. Introduction</b>	<b>49</b>
<b>2.2. Morality- naturalistically understood</b>	<b>50</b>

2.2.1. <i>Responsibility</i>	51
<b>2.3. Strawson and reactive attitudes</b>	<b>53</b>
<b>2.4. Revision instead of revolution</b>	<b>57</b>
<b>2.5. Moral luck versus ultimate equality</b>	<b>58</b>
<b>2.6. Criminal care</b>	<b>62</b>
<b>2.7. Conclusion of part two</b>	<b>65</b>
<b>3. Thesis conclusion</b>	<b>67</b>
<b>4. Sources</b>	<b>69</b>
4.1. <i>Books:</i>	69
4.2. <i>Articles and electronic sources</i>	70

## 0.0. Abstrakt på norsk

Denne masteroppgaven argumenterer for posisjonen *hard inkompatibilisme* i fri-viljedebatten. Er fri vilje og determinisme mulig i et determinert univers? Dette gamle filosofiske problemet er fortsatt svært aktuelt og debatten lever ennå i beste velgående.

Hvordan vi tar tak i dette problemet kommer an på hvilken definisjon av begrepet ”frihet” vi bruker og på om vi relaterer konseptene fri vilje og moralsk ansvar med hverandre eller ikke. Jeg argumenterer for at den posisjonen som heter *hard inkompatibilisme* er den posisjonen som er mest koherent med naturalisme, gitt den forståelsen av frihet jeg støtter. Med naturalisme, menes et synspunkt som ikke motsier naturvitenskapene. Jeg sikter på en forståelse av handling og moral som ikke introduserer unaturlige fenomen.

Mange argumenter for hard determinisme når det kommer til menneskelig handlingsevne er også tilfredsstillende koherente når det kommer til naturalisme. Men disse har svakheter som ikke rammer den harde inkompatibilismen. Indeterminisme er en like stor trussel mot frihet som determinisme, så den beste løsningen er å benekte frihet og heller fokusere på kontroll som det middelet vi har til å nå våre mål osv. Jeg forsøker å vise at med den rimeligste forståelsen av hva frihet er, er ikke determinismen kompatibel med fri vilje- i motsetning til hva kompatibilister argumenterer.

Problemstillingen min er: ”kan hard inkompatibilisme (med en naturalistisk redegjørelse av moral) tilby en løsning til problemet med fri vilje og moralsk ansvar, som ikke krever en revolusjon av våre reaktive holdninger?”

Forholdet mellom fri vilje og moralsk ansvar er det som står i veien for den kompatibilistiske løsningen; det virker urettferdig og urimelig å holde noen ansvarlig for noe de ikke kunne ha gjort annerledes, og for å være en person de ikke har designet. I denne oppgaven viser jeg at man må akseptere et sterkt konsept om frihet for å argumentere for at menneskelige agenter har moralsk ansvar for handlingene sine. Jeg vil dog, skille mellom *endelig moralsk ansvar* – som krever fri vilje, og *praktisk moralsk ansvar* - som jeg mener ikke krever fri vilje. Jeg vil forsøke å vise at evnen til å handle moralsk ansvarlig ikke har noe med frihet å gjøre, ved å prøve å gi en naturalistisk redegjørelse for hva moral kan være- eller i det minste ved å argumentere til fordel for å se moral som noe naturlig. Og at menneskelige agenter kan være praktisk moralsk ansvarlige selv om hard inkompatibilisme er sant. Men, de mangler endelig moralsk ansvar så det kan ikke rettferdiggjøres å fordømme mennesker for handlingene deres. På denne måten ender jeg opp med å benekte behovet for en revolusjon av våre reaktive holdninger som et resultat av en tro på hard inkompatibilisme. Muligens roer jeg også nye av frykten til de som tror at det å benekte fri vilje kan føre til kaos

når det kommer til moralsk opptreden. Jeg ender opp med å konkludere med at en tro på hard inkompatibilisme ikke fører til særlig forandring når det kommer til moralsk oppførsel. Hard inkompatibilisme-løsningen jeg foreslår gjør at vi kan beholde våre reaktive holdninger slik de er beskrevet av Strawson. Vi kan dømme når vi føler at noen virkelig kunne gjort noe bedre, men forskjellen mellom dette og den kompatibilistiske løsningen (på fri vilje problemet) er at grunnlaget for fordømmelse forsvinner på grunn av dets avhengighet av frihet.

### 0.1. Abstract in English

This thesis is an investigation into the free will debate and the problem of free will. Is free will and moral responsibility possible in a determinate universe? As old as this problem is, philosophers are still grappling with it and the debate on this issue is still very much alive.

How we address this problem depends on what definition of freedom we use, and it depends on whether we relate the concept of moral responsibility to freedom or not. I argue that the position called *hard incompatibilism* is the position that coheres best with naturalism given the definition of freedom I support. By naturalism, it is meant a view that doesn't contradict natural sciences. I aim for an understanding of action and morality that doesn't introduce any phenomena that isn't natural.

Many arguments in favor of hard determinism when it comes to human agency are also satisfactory coherent with naturalism, but has weaknesses hard incompatibilism avoids. Indeterminism poses just as much of a threat to free will as determinism does, thus the best solution is to deny freedom all together and rather focus on control as a means to achieve our goals etc. (to make things happen essentially). I attempt to show that with the most reasonable understanding of the concept of "freedom", determinism is not – as compatibilists argue – compatible with free will. My main question is "can hard incompatibilism (with a naturalistic account of morality) offer a solution to the free will and moral responsibility problem that does not require a revolution of our reactive attitudes?"

The relationship between free will and moral responsibility is what stands in the way of the compatibilist solution; it seems unfair and unreasonable to hold someone morally responsible for something they could not have done differently and for being someone they have not designed. In this thesis I show that one needs to accept a strong concept of freedom in order to argue that an agent is morally responsible for his actions. I will, however, distinguish between *ultimate moral responsibility* - which requires true free will, and the ability to

respond to moral situations as a moral agent; *practical moral responsibility*- which I claim doesn't require free will. By trying to give a naturalistic account of what morality might be- or at least by arguing in favor of seeing morality as something natural, I will try to show that the ability to act morally responsible has nothing to do with the ability to "do otherwise" or design oneself, and that human agents can therefore be practically morally responsible even if hard incompatibilism is true. However, they cannot justifiably be subject to condemnation for their actions because they lack ultimate moral responsibility. This way, I end up denying the need for a revolution of our reactive attitudes as a result of believing in hard incompatibilism. And, possibly also calming the fear of those who think that denying free will can lead to chaos, when it comes to moral conduct. I end up concluding that a belief in hard incompatibilism would leave us pretty much with status quo when it comes to moral demeanor. The hard incompatibilist solution I'm suggesting allows us to keep our reactive attitude patterns the way they are described by Strawson and to judge when we feel like someone could have done better. However, the difference between this and the compatibilist solution is that the grounds for condemnation disappear because of its dependence on free will.

## **0.2. Pretext**

My work with this thesis started with an interest in the concept of evil incarnated in certain individual humans, perhaps a strange springboard to the free will debate. The possibility for humans to be evil is something people may or may not believe in, but is often given credit as reason enough or almost enough for what I will call "enormously cruel acts" that the vast majority of humans could only do under extremely special circumstances- if at all. The way I saw it was that to be evil and deserving of condemnation, one has to have an evil will that is truly free and one's own.

I will look at various possible definitions of freedom, but for now it is enough to say that it seemed to me that calling someone evil was to cut the explanation for the enormously cruel acts very short. It also seemed to me that with closer examination and given some of the explanations that at first were hidden behind "evil", enormously cruel acts can sometimes make a lot of sense if looked at with the same tools we use for figuring out why other curious events take place in the world, namely by searching for various possible and probable causes for the effects we witness. The child soldier is an example of someone who can, in many people's eyes, often be relieved of all or almost all blame for their actions, given that people

know what it means to be a child soldier. The condemnation following the entitlement of “evil” would for many seem completely unfair in this case. When it comes to many other persons throughout history however, we often do call them evil (Adolf Hitler, Josef Stalin, Kim Il Sung, Idi Amin, etc.etc). I find it plausible that there should be causes for these people`s actions just like there are causes for the actions of the child soldier. This creates a dilemma. As most people and philosophers would agree, that while of course there are causes for every event in the world, they would also express that this cannot and should not excuse any evil person from the condemnation evil deserves. Yet it intuitively appears that if we accept or assume that human actions are determined then no one can really have the kind of freedom necessary for evil. This is how I came to be interested in the free will debate and its relation to the question of moral responsibility.

### **0.3. Introduction**

Hard incompatibilism denies human freedom and ultimate moral responsibility. The main question that I address in this thesis is: Can hard incompatibilism (with a naturalistic account of morality) offer a solution to the free will problem, or more precisely; the problem of moral responsibility, that doesn`t require a revolution of reactive attitudes?

The idea of a revolution of reactive attitudes suggest that if hard incompatibilism or hard determinism is true, then we can no longer blame or praise each other for the actions we do, and/ or this will lead to a view of humans as something mechanical with no self-control.

The discussion of my main question leads me to consider the following three “sub-questions”:

- 1) Why is a libertarian understanding of freedom preferable to a compatibilist understanding of freedom?
- 2) Why should moral responsibility have anything to do with freedom? –Why not hold that we do have freedom but not the kind of moral responsibility that make us condemnable?
- 3) How might the negation of freedom and of ultimate moral responsibility fit in with and possibly benefit society and the practicalities of dealing with “bad” (criminal, violent, mean, cruel, etc.) individuals?

The first two of these questions will be attempted answered in part one and the third in part two.



The free will debate is important because it deals with one of the capital questions of philosophy. And while still remaining unsolved, the problem of free will continues to be an ever-hot topic. Though the debate has stayed in its well-deserved highly prioritized position since antiquity, the questions involved should in my opinion receive increased interest from the philosophical as well as non-philosophical community as the levels of expertise in most sciences grow. This is especially so when it comes to the knowledge and theories from neuroscience, psychology and biology. As it stands with the status quo, one can ask two different philosophers the question “do humans have free will?” and get a positive reply from one and a negative one from the other. The same question can be directed toward two different physicists with the same result. Whether or not it lays in the cards of the natural sciences to answer all the questions about how humans act and for what reasons, the assignment for philosophy (in this context) is to formulate the right kind of questions. So we can avoid neuroscientists going off looking for free will on brain scans without really having a clear idea or opinion on what constitutes freedom of will. Regardless of there being a possibility for settling the quarrel once and for all, developments in fields like neurology will continue their impact on many aspects of human life. One of these aspects are justice systems, which are in most countries based on the assumption that humans are morally responsible for their actions and can meaningfully be imposed punishments for breaking laws etc. The line between evil and disability fluctuate with our levels of understanding.

It has been assumed or at least wide spread believed that freedom is essential for moral responsibility since the beginning of the western philosophical discussion. Thus the debate about whether or not humans have free will has most often also been a debate about moral responsibility, and many of the arguments to support the belief in free will have been motivated by a wish to “preserve” moral responsibility. What might have started out as epistemic curiosity is very often accompanied with a concern for the welfare of societies; a welfare many philosophers feel can be threatened by a belief in hard incompatibilism/hard determinism -or at least if one of those two positions turned out to be true. In my opinion, a philosophical debate should aim to discover truth; do we have freedom or not? Concerns about possible consequences of philosophical thoughts should be second order priorities. This opinion is echoed in the structure of my thesis; first I look into the debate on free will and arguments for the different positions, and then in part two I turn to consider the possible ethical implications of the position I hold.

The thesis is written, and should if possible be read, in the context of metaphysical naturalism; the thesis that the natural sciences are our best guide to what exists in the world

and that the world does not include any supernatural entities. While the term “naturalism” is so wide that almost no contemporary philosophers reject it, I find that emphasizing it is still useful as it sets a tone for the reading. Other directions of philosophy, such as existentialism would pose quite different premises for an investigation of evil, morality, freedom, human nature, etc. and is best seen as different board games with rules that differ greatly from the ones of the game I’m playing. I will therefore not spend any time or effort defending my arguments from the viewpoints of positions coming from vastly different territories of the philosophical landscape, but keep my business within the walls of a kind of metaphysical naturalism - to the best of my knowledge.

My thesis can be divided into two main parts. In the first part I argue in favor of the position called hard incompatibilism. This is where I try to answer the first “sub- question” mentioned above about why I find an incompatibilist definition of freedom preferable to a compatibilist one. Here I will try to show that we should not accept the compatibilist concept of freedom. The main reason for this is the compatibilist practice of, instead of saying something about what freedom is, they try to reduce it to something else- something that is not recognizable to us as “freedom”. That is based on our common- sense understanding of the concept “freedom”. I use the terms freedom and free will interchangeably.

In part one I will also assess the second “sub-question” which revolves around what reasons we have for relating freedom with moral responsibility in the first place. The first part starts off by presenting the free will debate and moves on to discussing various definitions of freedom as well as the concepts of control and the thesis of determinism. Then I look at the various positions in the free will debate, presenting the views and some problems with them and hopefully give a good account of why I favor hard incompatibilism.

In the second part I attempt to answer the third sub-question of my thesis. This concerns the question of what the moral implications of the position I hold in the free will debate could or should have for society were it to be commonly believed. I will then compare that to my impression of status quo. Hard incompatibilism is generally considered a negative approach to humanity, which should not be believed even if it is argued well for because it seems to imply that we are not morally responsible for our actions and this belief can be damaging for moral behavior. In this part, I suggest an alternative view that avoids the main problems that are often associated with denying freedom. I will try to explain why morality should be seen as something evolved; a character trait for the social and intelligent animal the human is. Thus I’ll attempt to show that freedom is not necessary for acting morally responsible or in

ways we find good. We act morally responsible because it is one of our evolved survival strategies.

As everything else in this thesis, how I portray morality is linked to what can and cannot reasonably be assumed about human nature, and our understanding of human nature should in my opinion not contradict what the sciences tells us. I'll try to deal with the terms as they come along and mainly put emphasis on what I take to be the general understanding of most terms. I make an exception for terms that I take to be ambiguous or when I need to give a term a particular hue for the purpose of this thesis.

The arguments I find advantageous are often arguments that go hand in hand with theory and evidence. In other words I'm writing this with a mindset filled with interest in and a significant degree of trust toward the sciences, and with the opinion that the kind of questions dealt with in here are more intriguing when philosophy can go somewhat hand in hand with science, or at least not completely distrust or divorce it. I aim to avoid being guilty of scientism. If this is understood as the practice of having a naïve and completely uncritical faith in the quality of the various studies being published hither and dither, science made pseudo by popular culture as well as the opinion that science should substitute philosophy and other arts, then I think I'm safe.

I find it important to emphasize that this is about considering what the world might be like, not to impose that this is the only good worldview or even correct one for that matter. Simply that among the contestants in the free will debate (in which many philosophers for good reason are not interested in participating), hard incompatibilism seems to me to have the best arguments. - One of the implications of believing that this position is true or likely is that one can stop believing in the possible existence for evil humans, which to me seems reasonable before even getting into the debate concerning free will and determinism (as mentioned in the pretext).

## **1. Part one: the free will debate and favoring the position of hard incompatibilism**

### **1.1. Introduction to part one**

In this part my main concern is to argue in favor of hard incompatibilism. I start with a short presentation of the free will debate and the main positions in it. Then I look at different ways of understanding the concept of freedom as a lot of the debate revolves around differences of

opinion in this matter. After that I discuss how and why freedom relates to moral responsibility. One section is devoted to the thesis of determinism. This is followed by the remaining sections that present arguments for and against the position of libertarianism, compatibilism, hard determinism and finally hard incompatibilism.

## **1.2. The free will debate**

Today, there are four main views that dominate the ongoing free will debate. First there are the libertarians who believe that determinism and free will is incompatible and for various reasons they lay their eggs in the free will basket and deny the determinism. They generally argue for their view by appealing to our intuitions, for instance the intuition that we are free. This is often because they value moral responsibility very highly and believe it is dependent on free will. A Libertarian may argue that her view is compatible with naturalism by appealing to indeterminacy in quantum mechanics. Central traditional libertarians are Immanuel Kant and Rene Descartes, and some contemporary libertarians are Timothy O'Conner, Peter Van Inwagen and Robert Kane.

The second group of contenders in the free will debate is the compatibilists. These philosophers hold that there is no problem with determinism and freedom both being the case and they offer many creative solutions to how this may be, most of which are definitions of freedom that might not conflict with the thesis of determinism. Some compatibilist views I include are the ones of Daniel Dennett and Harry Frankfurt.

Hard determinism is the third position I'm going to give attention to. Hard determinists agree with libertarians that free will and determinism is incompatible. They place their bet on determinism instead of free will, also for various reasons that will be elaborated further in the section devoted to hard determinism. This position is often seen as a negative and sometimes even cynical and hopeless point of view; among other reasons because hard determinism seem to "free" (pun unintended) people from their moral responsibilities. But also because it may encourage a defeated attitude and a belief that one cannot help oneself. Well-known hard determinists are Sam Harris and Ted Honderich.

Hard incompatibilism is the fourth and final view I will include in this thesis and the one I wish to defend. The philosophers holding this position generally argue that because of the nature of freedom and the nature of humans, freedom is something that humans cannot have, regardless of determinism being true or not. Like compatibilists they agree that indeterminism would allow for no more freedom than determinism, but they do not concern

themselves with finding loopholes in definitions or metaphysics. Unlike hard determinists they also do not commit so fully to determinism being true. The hard incompatibilist I focus on is Derk Pereboom.

### **1.3. Freedom**

#### **1.3.1. Different philosophical understandings**

When philosophers discuss whether or not we can have free will in a deterministic universe, their understanding of freedom is central to their arguments. Their definitions of the concept of “freedom can differ enormously and it is central to look at which concepts that are at play in order to evaluate the debate. Following is a description of two of the most common types of understandings of what freedom is.

##### **1.3.1.1. The garden of forking paths.**

Traditionally it has been common to understand “freedom” as the ability to do otherwise. This understanding of freedom is libertarian and libertarians generally base this ability on indeterminism. The garden of forking paths is a metaphor for a human life filled with choices, the past is closed and everyone agrees we can no longer do anything about that, but the future is open and holds various alternative outcomes. It is entirely up to the agent which path he chooses and thus what his future will be like. The ability to choose otherwise means that given all other things equal, the agent won't necessarily make the same choice on Thursday as he did on Wednesday. This is commonly one of the first steps of nitty gritty freedom investigation people do when they get in to the free will problem. “Can I choose randomly between an apple or a pear or is that too predetermined by antecedent events? - Can I choose to scream out loud in the cafeteria out of the blue?”. At first it seems absurd to suppose anything other than “of course”, however, if determinism is true then we cannot do things out of the blue or randomly. If determinism is true, we do not have the ability to do otherwise.

##### **1.3.1.2. The source model of control.**

Some philosophers have suggested that we should not understand “freedom” as the ability to do otherwise. This is because they think such a definition is too restrictive. According to the source model of control an agent is free if he can be the source of the causes that lead him to make one choice over another. He does not need the ability to do otherwise and in fact the ability to do otherwise might even render him less free than he is without that opportunity.

Strictness about how an agent is the source of his actions varies according to which philosopher is asked, some feel it is enough to state for instance that the agents interests are part of him and cause some of his actions, while others will always move beyond that and ask “but where did the interest come from?”- this is a practice that can continue until one is well outside of the scope of the individual’s agency.

A libertarian concept of freedom that fits in with the source model of control rather than the ability to do otherwise is agent-causation. This is the idea that a free agent can cause his own actions in a special way, independent of prior events.

According to many compatibilists such as Frankfurt and Dennett, an agent’s ability to act according to his character and values qualifies as all the freedom we should want and is sufficient for moral responsibility. Regardless of why an agent has acquired the particular set of values he has or why he has developed into the person he is, he is free and morally accountable.

### 1.3.1.3 Kane’s five freedoms

The two understandings of freedom I have described above appeal to different concepts of freedom in the free will debate. Within these two concepts, there are several versions of each, particularly the source model of control. As a conclusion to his book *A Contemporary Introduction to Free Will*, Kane sum up five different freedoms.

The first one he mentions is the *freedom of self-realization*. This is the idea that you are free if you can realize your wants and wishes without outer restrictions, pretty much what is acquired through political and religious freedom, and human rights.

The next freedom in Kane’s summary is *freedom of (reflective) self-control*, in essence the first kind of freedom plus freedom from inner obstacles like mental illness, poor self-control, addiction etc. Advocates of this kind of freedom are Harry Frankfurt, Plato and Aristotle. According to Frankfurt, one can be morally responsible with reflective self-control even though one could not have done otherwise. Understanding the difference between right and wrong and having enough self-control to act consistently with this understanding renders a person morally responsible.

Third, Kane introduces *freedom of self-perfection*. This freedom is much like the freedom of (reflective) self control- where one acts according to reasons etc., only with this kind of freedom, the agent not only acts according to his reflected reasons and understanding of right

and wrong, but additionally he knows what the *right* reasons are. Kane mentions Susan Wolf as the one advocate of this kind of freedom. She gives an example of a Kim Jong Il like dictator's son to describe a person who is not coerced, mentally ill or a wanton (someone who follows his desires uncritically), and who does reflect on who he wants to be and what his values are and still is not free because his values are wrong is. In order to be truly free, one needs the ability to perfect and correct oneself. According to Wolf this is perfectly possible, unlike self-creation.

Freedoms number four and five are the *freedom of self-determination* and the *freedom of self-formation*. Contrasting with the previous three freedoms, these freedoms do not appeal to compatibilists. These are libertarian kinds of freedom and the freedom of self-determination consists of “the power or ability to act *of your own free will* in the sense of a will (character, motives and purposes) of your own making—a will that you yourself, to some degree, were *ultimately responsible* for forming.” (Kane, 2005, p. 172). To have freedom of self-determination one must also have “*The Freedom of Self-formation*: the power to form one's own will in a manner that is undetermined by one's past by virtue of *will-setting* or *self-forming* actions (SFAs) over which one has plural voluntary control.” (Kane, 2005, p.172). As long as an agent have self-formed, he does not need to have the ability to do otherwise in the garden of forking paths when it comes to the self-determining acts of everyday life. He can however take full credit for the self-forming as this is an activity separated from determinism.

### 1.3.2. Common-sense

In the free will debate, it is often a problem that both compatibilists and the various incompatibilists in the free will debate claim to have “most peoples intuitions” or commonsense on their side. That obviously cannot be the case though it might be the case that people differ in their pre philosophical intuitions about freedom and determinism.

Eddy Nahmias, Stephen Morris, Thomas Nadelhoffer, and Jason Turner have conducted a study to try and tackle this problem. The participants in their study were undergraduate philosophy students who had yet to study the free will problem. Arguably, the choice of participants was a huge mistake. Philosophy students regardless of having studied the free will problem will most likely be familiar with it anyways, and I also find it probable that they would have an idea of their own position colored by philosophical insights and ideas non philosophers would be unfamiliar to. It seems like the best way to get hold of what the

common- sense intuitions about freedom and determinism are, would be to ask people without a particular interest in philosophy. That's why I think the conclusion of Shaun Nichols paper "The Folk Psychology of Free Will: Fits and Starts" that he based on studies of small children is preferable to the one of Nahmias and his crew. Interestingly, the two conclusions suggest pretty much the same. According to Nichols "...it's possible that folk psychology, construed broadly, is committed both to agent-causation and to a deterministic story about psychological processes." (Nichols, 2004: 499) He claims that this is an inconsistent metaphysics of agency. Nahmias and crews interpretation of their own study concludes that folk psychological intuitions are compatibilist. They make this claim because of the co existence of the two intuitions of freedom and determinism in most people or commonsense. However, the kind of freedom commonsense finds compatible with determinism is agent-causation: a libertarian kind of freedom. For this reason I think it is legitimate of Nichols to describe it as an inconsistent metaphysics of freedom. Because folk psychology or common-sense is not philosophy and doesn't need to concern itself so much with the consistency of intuitions it can combine libertarian freedom with determinism without worries. But because philosophers can't do that, arguably one should be skeptical toward the claim that the common-sense understanding of freedom and determinism is compatibilist. A compatibilist account of freedom is very different from a libertarian one, which is what the common-sense understanding of freedom mostly, resembles.

According to the above-mentioned studies, common sense tells us that we have free will in the sense of agent causation- the ability to create causes for our actions. Because agent causation seems for many (non libertarian philosophers at least,) to be unintelligible and to conflict with the thesis of determinism, various alternative proposals to what freedom may be have come about. Such suggestions can be that freedom is to act for ones own reasons, regardless of those reasons being a part of a causal chain that starts way before and outside of the agent. This is the general compatibilist understanding of freedom.

When it comes to justifying moral responsibility, there is in practice a massive difference between the folk psychology understanding or intuition of freedom and the general compatibilist one. In my opinion, the term "freedom" should be reserved for the expectations of common sense; it should mean the ability to do otherwise and / or to create ones cause. The reason for this is that we have good words to describe what compatibilists call freedom, for instance control, self-control, self-discipline etc. When it comes to the kinds of freedoms



discussed in the free will debate, as apposed to political freedom, most of us are talking about something which can provide us with moral responsibility and a kind of elevated self ownership. If freedom exists it should grant us the true ability to act freely and unchained by causal relations in nature. In this context it has an almost magical ring to it. And while libertarian understandings of freedom are being rejected by compatibilists because they seem mysterious and unintelligible, they seem to me to be that way because they would not work if they were different. The way the definition of God (including adjectives like omnipotent, omniscient, eternally good, the first uncaused cause etc.) has to be something grand and vague for it to serve justice to the concept, freedom too has to appeal to this desire of something supernatural or at least mysterious for it to be justified as something more than just control. Whether people can believe in it is an entirely different case. Regardless of our believes in the existence of freedom, in order to take the word to have this meaning, one has to either to adopt a libertarian view in the free will debate, or one of the two positions that denies that humans have freedom. It could be tempting to just stop there and perhaps I could if I was only concerned with the question of freedom. It is however interesting to investigate how compatibilists defend the possibility for moral responsibility thorough various definitions of freedom.

### **1.3.3. Intuition and experience**

Based on the two sources of investigation, I have now given an account of what the common sense understanding of freedom seems to be. Left for discussion is why most of us seem to have this intuition. Everyone who is a person has access to the answer though, which is our experience of freedom. Even without defining what freedom should or could be metaphysically speaking, most of us feel strongly that we have it. In many ways libertarianism seems like the more rational choice when forced to choose between two opposing philosophical thesis's about the world; this is the one we have the most intimate first person relationship to. Even hard incompatibilists, who believe that freedom is something humans cannot possibly have regardless of the universe being deterministic or indeterministic, have strong feelings of freedom. When we make a choice, it seems like we could truly have chosen another alternative and when we make bad choices we feel stupid, regretful and sometimes guilty about it. Hard incompatibilists get these feelings just like everybody else, even though they believe that in order for them to have acted differently,

something in the choice situation must have been different (their own wishes for instance) and thus they end up denying the idea of freedom.

Without drawing the hard incompatibilist conclusion however, there seems to be other reasons to question our experienced based belief that we are free. In the consciousness area of philosophy of mind and perception there is a lot of interesting discussions revolving our introspection; our ability to know facts about ourselves through the inward gazes of which Descartes made the fundament for all knowledge.

There are reasons to think introspection is not as reliable as we have assumed. Eric Schwitzgebel writes in his paper “The Unreliability of Naïve Introspection” that “the introspection of current conscious experience is both (i) possible, important, necessary for a full life, and central to the development of a full scientific understanding of the mind, and (ii) highly untrustworthy.” (Schwitzgebel, 2008). What he means by this is that “We fail not just in assessing the causes of our mental states or the processes underwriting them; and not just in our judgments about nonphenomenal mental states like traits, motives, and skills; and not only when we are distracted, or passionate, or inattentive, or self-deceived, or pathologically deluded, or when we’re reflecting about minor matters, or about the past, or only for a moment, or where fine discrimination is required.” (Schwitzgebel, 2008).

When we introspect our will as free I suspect it is because we fail to see the causes of the will being the way it is. The libertarian might assume there are no causes for the will and this might be based on the introspection where they see only the desire for the will, not the cause of the desire. (Though, I suppose we quite often do introspect the causes of our will; “I was brought up in a Christian home so I want to pray every night before I go to bed etc). Assuming that there are no causes for an event simply because you can’t see the causes is very naïve.

Using what Schwitzgebel says about introspection as an argument in favor of incompatibilism can of course backfire. That is also assuming something without having a satisfying amount of factual reason to make that assumption and this is also quite naïve. According to Schwitzgebel, Descartes was wrong when he claimed we have a greater access to our own minds than the world around us. The opposite is more likely to be the case based on the unclarity of our own introspections.

The first scientist to prove that decisions on how to act can be decided subconsciously

while we experience it as being consciously made was Benjamin Libet. It is likely that in many (or most) cases of decision making, this is happening; the decision is made unconsciously based on whatever reasons the brain has to favor that decision and the conscious *you* only feels like it is “my will” or “the will of my consciousness”, and has an experience or mastery over the action performed as a result of the decision making. In reality it was just following orders from “below” from subconscious reflections or calculations being done by the rest of the brain. The experiment Libet describes in his paper “Do We Have Free Will” was conducted while the subjects focused on an oscilloscope ‘clock’; a clock in which a spot of light revolves around the center once every 2.56, pausing at each point for 43 milliseconds (instead of the normal one second pause of the second hand on a regular clock). The subjects were given the task of flipping their wrists whenever they wanted, but were asked to notice at which point in time on the oscilloscope they became aware of the decision to act. The experiment showed that regardless of the subject experiencing the wrist flick as preplanned or spontaneous, the brain activity leading up to the action started significantly prior in time to the subjects becoming aware of making a decision. (Libet, 1999, p. 50). Interestingly, Libet also found that while the subjects seem to simply be informed by their sub consciousness that “I’m going to act now”, they were able to veto this unconscious decision-making consciously. This led Libet to conclude that we do have free will, in the form of a free “won’t”.

I do not doubt Schwitzgebel being right about our ability to introspect being generally much worse than we think, but it must be just a bit more functional than dysfunctional. We do succeed in communicating to others how we feel a lot of the time and that must mean that we must have a certain clue about our inner life. It seems likely that introspection is a tad more functional than dysfunctional for the simple reason that we keep using it- that it has survived the natural selection. Perhaps it is like the case with intuition, it works to a satisfactory degree most of the time and it is few occasions when it doesn’t work that is interesting to this thesis. Introspection of a free will is a misinterpretation of what might actually be there for us to introspect: a significant degree of control in many situations.

### **1.3.4 Control**

So, having established that the term freedom should remain loyal to the folk psychological concept of freedom and that this concept is a libertarian concept of freedom, or refers to a

libertarian type of freedom, both compatibilists and hard determinists/ hard incompatibilists will have a hard time making sense of our introspection of freedom being a true testament to our libertarian freedom. Though introspection is interesting when it comes to a lot of questions in philosophy of mind, in this case I don't think the introspection is the problem, but rather the interpretation of the introspection. In other words, if we interpret our experience of following our own wills and consciously deciding how and when to act as free will then that might be a false conclusion. I think the reasons why most of us believe we have free will is because we introspect that we can in many situation do what we want, and then we mix that with our folk psychological concept of freedom. But it can be even more severe than that; we might even be wrong when we introspect that we are doing what we want.

We may not have a libertarian free will but it seems obvious that we very often have the ability to get what we want, bring about conditions we want and desire and generally make things happen. None of this is conflicting with us being a part of a larger causal chain of events and causes. The general compatibilist claim is that our choices can be free even though we could not have done otherwise or made our reasons/ causes for making the choice because they are based on reasons we identify as our own, in according with our deep characteristic values etc. I agree that we have what the compatibilists describe here, but not that it is freedom (because of the factors that determine what we want). Rather it seems much more appropriate to call it control. Thus, compatibilists and hard incompatibilists/determinists can have the same or a very similar understanding of agency but will continue to differ in the matters of moral responsibility.

Making use of the term control where compatibilists would say we have freedom provides three benefits; one, that autonomy and the difference between voluntariness and involuntariness/ outside force etc., is appreciated. Two; we can avoid violating the commonsense expectations of what freedom means. And three, we can avoid ascribing *ultimate moral responsibility* to agents for actions they could not have done otherwise or have been the source of.

Sometimes control can be used as a kind of synonym to free will. That is obviously not what I want to do in this thesis, but the two are related. I view free will as *ultimate* control, and control as what allows us to often make our wills happen.

#### 1.4. The relationship between freedom and moral responsibility

If the free will debate really was just about the concepts of freedom and determinism and whether or not those were compatible, then a compatibilist style definition of freedom wouldn't sound as problematic. It still wouldn't sound good to me because I do have the idea stuck in my head that freedom should "free us" from some of the attached strings and provide us with a magical value that sets us apart from other animals.

The serious problem in the free will debate arises when it becomes evident that people/philosophers attach moral responsibility to freedom. When that happens (and it has of course been happening from the antique beginning), the compatibilists versions of freedom become much more problematic. To say that a person is free because he follows his own will in life sounds reasonable enough, but when the same person does a bad thing that he could not have done otherwise (in accord with compatibilist beliefs) it doesn't necessarily make sense to hold him ultimately morally responsible for his actions. It can make sense to hold people morally responsible for practical reasons when they follow their own will, but to claim that they are ultimately morally responsible is to take things too far. One good objection to this could be to answer this with "too far where? Is it ever not practical and pragmatic to hold someone morally responsible for doing what they want?". If causality works the way I've understood it, then if I were born Hitler and Hitler me, I would do what he did and he would live my life exactly as I live it now. And if that is true and come the Christian or Islamic judgment day, he would be extremely unlucky and I would be lucky (or at least presumably luckier than Hitler). And that doesn't make any sense.

I'll also get back to moral luck in part two. I'm trying to show that we cannot have ultimate moral responsible for our actions, and the occurrence of moral luck seems to show that it is sometimes up to luck who gets considered blame or praiseworthy by the theories that support that humans are ultimately morally responsible for at least some of their actions. On the one hand, this way of reasoning about freedom and moral responsibility is completely ridiculous because most of us probably do not believe that it ever comes to the "ultimate point" of a judgment day. On the other hand however, we do seem to judge people way beyond the sentence they are given in the justice system.

Freedom and moral responsibility are concepts often viewed by philosophers and non-philosophers as two sides of the same coin; you can't have one without the other. As we will see, what it takes for an action to be free depends on which understanding of freedom we endorse. As mentioned, in my thesis I find it meaningful to separate between *practical moral*

*responsibility*, and *ultimate moral responsibility*. This practice has been rightfully criticized, as I will show in part two, however, there is in my opinion a good point to making this separation. If we view morality and responsibility through naturalist glasses, as abilities we have because they have proven useful and survived natural selection, it does not immediately seem necessary to have free will in order to explain them. Ultimate moral responsibility on the other hand seem to indeed require freedom because for something to truly be our fault, it seems necessary that we have really caused it to happen. If freedom does not exist, then ultimate moral responsibility is impossible, but our abilities to act morally and take responsibility are not affected.

## **1.5. The idea of determinism**

Being a thesis that centers round the idea of cause and effect, determinism concludes that all events, both past, current, and future, are pre decided by antecedent events that go back to the big bang or probably further.

### **1.5.1. Confusions about determinism**

As the Stanford Encyclopedia of Philosophy states, it is a common mistake to confuse determinism with predictability. The mathematician/ astronomer Pierre-Simon Marquis de Laplace came up with the thought experiment known as “Laplace demon”. “We may regard the present state of the universe as the effect of its past and the cause of its future. An intellect which at a certain moment would know all forces that set nature in motion, and all positions of all items of which nature is composed, if this intellect were also vast enough to submit these data to analysis, it would embrace in a single formula the movements of the greatest bodies of the universe and those of the tiniest atom; for such an intellect nothing would be uncertain and the future just like the past would be present before its eyes. (*Essai philosophique sur les probabilités*, 1814)”(Information philosopher). It is tempting to agree with Laplace in this but as Carl Hoefer states in the SEP article “.19<sup>th</sup> and 20<sup>th</sup> century mathematical studies have shown convincingly that neither a finite, nor an infinite but embedded-in-the-world intelligence can have the computing power necessary to predict the actual future, in any world remotely like ours.” (Hoefer, 2010). First of all, this doesn't put a stopper to the thought experiment of Laplace demon (; the demon could be infinite and not embedded in the world), and secondly it does not pose a threat to determinism as a theory. As

clarified earlier: the definition of determinism I'm using in this thesis is: a theory that suggests that all events have causes and that the causes they have are sufficient to make them happen. This way it seems that the world is a closed system of causal relations, chains of cause and effect and nothing physical stands outside of this system. Claiming determinism for true or likely is not the same as claiming predictability as probable or even possible, though one might still hold Laplace demon as logically possible. Heisenberg's uncertainty principle may be another thing worth mentioning as an obstacle for predictability but not necessarily for determinism.

According to determinism the future is not open though it seems that way to us because we have no way of knowing what it holds. This ignorance can be said to be bliss for humans; we try our best and this can often lead to good outcomes for us. When someone has died, determinists can say that he or she could not have lived their lives differently. However, until the point of death, the factors that can change a human's life course are almost unlimited, so we do not need to feel "trapped" even though we hold determinism to be likely or true.

The thesis of determinism is also often confused with fatalism. The confusion is understandable and caused by the fact that whatever has been predetermined to happen will happen both according to determinism and fatalism. The significant difference however is that in a deterministic world, the events that happen do so because of predetermined events that caused the now happening events. In a fatalistic world your fate will happen regardless of what comes before your fatal events set in. You may try to run away from your fate but will see it happen regardless of the level of your ingenious efforts to try and escape.

### **1.5.2. Most philosophers endorse determinism**

Both compatibilists and hard determinists endorse determinism. Hard incompatibilists might find determinism likely, but do not have to commit to the universality of the idea and libertarians deny it. Still depending a bit on whom you ask, determinism has been refuted by science. This is basically because of randomness of particle movements in quantum physics/mechanics and the way radioactive substances decay: indetermined. Logically it must be the case that if even just one event in the history of the universe is indeterministic, then the universe is indeterministic. A hard determinist can reply to the quantum mechanics argument by suggesting that the particles do not move around randomly and undetermined, but that it simply appears that way because science hasn't figured out the system yet. It is in my impression however more common to just draw a line between micro and macro levels of

physical objects and claim that because some things may be undetermined on a micro level, that does not mean that determinism cannot be true for the movement of macro level objects like humans.

While it may seem strange to believe in determinism when most of scientists claim it has been proven faulty, the idea of more or less necessary causality is a very common intuition. Among philosophers at least, it might even be just as strong as the freedom intuition. Based on the numbers from the Chalmers and Bourget survey, 59,1 % of philosophers asked informed that they are compatibilists and 12,2 % hold a no freedom position. Of these there must be some hard determinists so maybe its reasonable to assume that about 65% of the philosophers asked believe in determinism. Now, is this because they simply don` t pay attention to the world of science or is it because they are not satisfied with the quantum mechanics denial of determinism? I would bet on the latter but that is not too important for now. What is of importance is that most philosophers (asked by Chalmers and Bourget) believe in causal necessity (57,1% believe in non-Humean natural laws). As Robert Kane points out, even though modern physics move away from the thesis of determinism, other sciences do the opposite. Psychology, neuroscience, biochemistry etc. offer explanations to our behavior that we are not aware of by causes outside of our control.

## **1.6. Libertarianism in general**

My conclusion about free will, one genuinely free in the non-determined sense, is then that its existence is at least as good, if not a better, scientific option than is its denial by determinist theory. Given the speculative nature of both determinist and non-determinist theories, why not adopt the view that we do have free will (until some real contradictory evidence may appear, if it ever does). Such a view would at least allow us to proceed in a way that accepts and accommodates our own deep feeling that we do have free will. We would not need to view ourselves as machines that act in a manner completely controlled by the known physical laws. (Libet, 1999, p. 56-57)

Libertarians in the free will debate must be separated from libertarianism in political contexts. What the two have in common is a concern for freedom and freedoms, but the libertarians of the free will debate want to investigate and argue in favor of the claim that at



least some human beings have free will at least some of the time. Libertarians in politics are concerned with individual political freedoms like that of speech, religion, economical freedoms, human rights etc. The same person can be a political liberal and a compatibilist in the free will debate.

A general libertarian concept of freedom is that it enables a person to do otherwise: to have a true choice between *a* and *b* for instance, unaffected by determinism. This is the garden of forking paths model of freedom. Libertarians believe that freedom is incompatible with determinism and that because freedom exists, the universe must have some degree of indeterminism about it. Some libertarians (Descartes among others,) have been substance dualists- they have accepted that determinism might be the case for nature and for the human body, which is natural, but that there is an aspect to humans that is more than natural. This is usually the mind or the soul and is of another substance than the rest of the things in nature. A general problem for philosophers who try and solve the free will problem through dualism is to explain how something non-physical or physical in a radically different sense can influence a physical body and thus cause “free actions”. Today, there aren't that many substance dualists left in the world of philosophy but there are however property dualists- I will get back to those later on. For now I will focus on the best reasons for being a libertarian in the free will debate today.

The first and seemingly foremost reason to believe in the kind of free will libertarians hold to be true is the feeling or intuition we have about freedom. Unless we are being forced, brain washed, tricked, coerced, controlled by other agents somehow or influenced by mental illness, drugs etc., most of us feel like we are free to do either *a* or *b* when *a* and *b* are quite similar in significant ways (like effort requiring etc.). For instance, it seems entirely up to me whether I buy an apple or a pear at the store. As long as the choice between *a* and *b* is sufficiently trivial, it seems that I could go back in time and not necessarily make the same choice in the same situation. If I have a reason that makes me prefer the pear to the apple for example, it is very likely that I will make the same choice if I find myself in the same situation again. According to the general libertarian I can however sometimes choose without reason. When I plan for or dream about my future, what it will hold appears to my imagination as open and undecided. Agent-causation, which is more in tuned with the source model of control is the idea that a person as an agent can “create” causes for action for instance through rational deliberation, careful consideration etc.

The second reason, and to many perhaps equally important as the first, is related to moral responsibility. To a libertarian, free will is a necessary trait to having moral responsibility- if

an agent does not have the ability to do otherwise it seems logical that he cannot be held morally responsible for his actions. Most libertarians value and believe in freedom for itself and not as a means to an end, but others have argued that we must believe in free will to preserve the sense of moral responsibility most of us have. This is of utmost importance as it can be reasonable to assume that society and civilization depends upon it. Philosophers who think along these lines are most likely haunted by fantasies about Hobbes natural state where everyone lives for himself and must fear the knife of his neighbor all his brutish and short life, should they be proven wrong about freedom and moral responsibility. (See for example Duus-Otterström, 2008) A libertarian would most likely be against the attempt to divide the ability to act morally and take responsibility, and having ultimate moral responsibility for ones actions, and I agree with them that being fully responsible for ones actions require that one has either the freedom to do otherwise or significant control over the source of ones actions, the source model of control -or preferably both.

The third reason for positioning oneself on the libertarian team in the free will debate is scientific findings of indeterminism in nature. Examples of this are the motions of particles described by quantum physics and radioactive decay. This may seem like the best reason to avoid the whole problem of free will and determinism, why assume determinism and compromise freedom when science has shown us that the universe is not in fact deterministic?

### **1.6.1. Some problems with general libertarianism**

When it comes to the first reason for believing in libertarianism it is not necessarily a problem that we have a feeling or intuition of freedom, but it is a problem to assume that freedom is a fact based on this intuition. Many of us also have an intuition about causality, and as libertarians would agree; that conflicts with a free will in the sense they are talking about. On the macro level, determinism seems to correct- physical theories on the macro level are deterministic – e.g. the theory of relativity.

Secondly, there are neurological findings that imply that we should not trust our introspection too naively. Experiments like those of Benjamin Libet have shown that what happens in the brain is not always what our conscious selves experiences. The unreliability of introspection can of course work against the “causality-intuition” as well, but this intuition has the backup of physical theories the macro level.

As for the second reason to hold libertarianism to be true, namely for the sake of preserving moral responsibility, there are ways in which we may think we do have moral responsibility without the freedom to do otherwise, or that we have freedom to do otherwise, or another kind of freedom but not moral responsibility. This is the claim of the general compatibilist and I will get back to this in my descriptions of compatibilism and hard determinism.

When it comes to the strategy of denying determinism about human actions and thus removing the tension between that and free will the main problem is to show in what way indeterminism can allow for any more freedom than determinism. If our choices are not determined by our reasons and reflections but may differ from what those would incline us toward, then chance or some other factor outside of our control could be the cause of our choice and how would that give us any more freedom than regular determinism? Say that John for example really wants to buy a new car. He has the choice to buy a motorcycle instead of a car, but he would greatly prefer the car. He needs it to get to work and he is also very fond of cars and hasn't owned one in a while so he is very excited about getting one. He has enough money to buy a car and the license to drive it. Cars are also completely available for purchase in the place where John lives. Apparently there are many causes for why John should end up buying a car. However, if indeterminism is true, then these causes may be present and even though nothing is really standing in John's way of buying a car (like him becoming acutely ill or something), he may actually end up with a motorcycle instead. If this were to happen then that would indicate that John had less control of getting what he wanted that he would have in a determined universe.

It is certainly interesting how most philosophers commit to determinism (59.1% are compatibilists according to the survey performed by Chalmers and Bourget, 2013) when science seems to deny it, however in the project of conjuring up an explanation of free will indeterminism seems to do us no more favors than determinism. Though some libertarian arguments may be better than others, libertarians are nevertheless faced with the difficulty of explaining how indeterminism could provide freedom without sounding mysterious or unintelligible.

### **1.6.2. The best account of libertarianism**

What leads to a belief in determinism is the assumption that events in nature are caused by antecedent events and that this causality is somehow necessary. This means that if what it takes for a candle to be lit on Wednesday is the candle, a lighter and someone to perform the task, then given that all three factors are functioning and all other things are equal- the candle

will be lit the same way on Thursday. In real life there are almost always too many variables to enable us to predict how real events will occur in great detail. Many events are however predictable enough so that we can end up with a belief in determinism based on our experiences of the world.

Arguably, the best argument for libertarianism is the claim that causes do not have necessary effects on events, but rather that certain causes tend to lead to certain effects. The idea is that the tendencies are strong enough for events to seem necessary or law based and somewhat predictable, but there is room for indeterminism and free will because of that lack of necessity between cause and effect. In thought experiments like the one with the candle above, or in real experiments where one isolates causes and effects from the rest of the world the results we get are artificial and will only really give us knowledge about tendencies, not actual causes because events that are not isolated in this way will always be under threat of extra factors. For instance, if we put a heater in an empty room with controlled ventilation and insulation, then the heater will cause the room to be at a predicted temperature. However, if we buy a heater and put it in our real life apartment there will be a lot of extra factors that cause the temperature to differ from what we set the heater to. Drafty windows, poor insulation, the temperature outside that day etc. can all be factors to make our apartment colder and factors such as many people in the room, sunny weather, cooking something on the stove etc. can make the temperature warmer than what we set the heater to. The heater will be a cause that tends to make the room warmer, but other factors may be present and cause a tendency to make the opposite happen. Simple, isolated events like those in experiments are not like events in nature and all we can know about causes is that they tend to lead to certain effects, nothing more. In their book, *Getting Causes from Powers*, Stephen Mumford and Rani Anjum argue that causation is powers manifesting (of the tendencies). This is based on a metaphysics of powers, something I will not further explore in this thesis but I find it interesting to consider the possibility that causes could only tend to effect events the way I think they necessarily do.

Regardless of me understanding metaphysics of powers, when it comes to providing us with a loophole for free will in nature, this indeterminism based on tendencies instead of necessity seems to 1; not really solve the problem of indeterminism (how do you get free will if your causes for actions also only tend to provide the wanted outcome?) like the rest of the accounts of libertarianism, and 2; to play on the weakness of human predictability as an argument against determinism. The fact that some events entail so many different factors that

it makes them difficult or even impossible for us to predict do not serve as a good reason to let go of our assumptions about laws of nature.

## 1.7. Consciousness, subjectivity and the free will debate

A couple of topics that very often come up in discussions about freedom and determinism concern consciousness and subjectivity. For a defense of libertarianism or a discussion of the possibility for a libertarian free will, the following may be very important. Libertarians may appeal to consciousness and subjectivity to explain or find room for their kind of freedom.

### 1.7.1. Consciousness

Kane mentions Roderick Chisholm as a well-known defender of agent-causation. This is as Kane states an attempt to solve the libertarian dilemma; that a free action cannot be caused by antecedent events in accord with determinism, but it also cannot be a product of chance or be completely uncaused. He calls the thing that explains this kind of freedom the “extra factor”：“...the extra factor is a special or unique kind of causal *relation* between an agent and an action that is not reducible to, and cannot be fully explained in terms of, the usual kinds of causation by events, occurrences, and states of affairs, either physical *or* mental.” (Kane, 2005, p. 45). It seems that defenders of agent causation are in need of an explanation of how agent causes can be such beginnings of causal chains as is claimed and of how agents can be first unmoved movers. They may go looking for this in the study of consciousness.

The experiments Benjamin Libet started to conduct in the 60ies and similar experiments and investigations of the neuroscience of free will are at best showing that decision-making sometimes starts before the conscious agent becomes aware of it- and experiences making the decision as he becomes aware of it. Protests against the validity of using these findings to say something about our free will or lack of which are usually based on a reaction of the simplicity in the decisions being made while conducting the experiment (for instance; to push a button at a given time etc.) and that if the decisions were to be more significant for the agent, perhaps he would become aware of it before it had already been made. It is tempting to use indications made from the results of experiments done by scientists like Benjamin Libet to conclude that the human will is subconscious, and therefore not free. Libet did claim however, that even though we seem to not have a free will, we do have a free “won’t”, our

consciousness can veto actions started subconsciously. This unconscious veto provides according to Libet a true free will that allows us to act in spite of known physical laws. “..But we have not answered the question of whether our consciously willed acts are fully determined by natural laws that govern the activities of nerve cells in the brain, or whether acts and the conscious decisions to perform them can proceed to some degree independently of natural determinism.” (Libet, 1999, p.55) While acknowledging that he has not proven this with his experiments, Libet suggests that we believe in a free will that is not compatible with determinism until it can be proven that we do not have it. He doesn't provide a suggestion of how our free will may go about this breaking of the rules of everything. It seems to me that if we go looking for free will in the brain, the only thing we may find is freedom according to a compatibilist definition, or what I would call control. So unless one believes like Libet that we might have a freedom that is set apart from the rest of nature, then understanding how decision-making relates to consciousness is irrelevant for the free will debate between compatibilists and hard incompatibilists/ hard determinists.

### 1.7.2. Subjectivity

Subjectivity is also a topic that keeps occurring when discussing free will. This can be another place where libertarians can hope to find the kind of free will they want and believe in. For anyone who accepts or bets on determinism at least when it comes to human actions, this is not significant, but it is in my opinion one of the strongest appeals to a libertarianism that is ready to deny naturalism. According to the survey conducted by Chalmers and Bourget (2013), 49,8% of philosophers believe in naturalism when it comes to metaphilosophy and 56,5% hold physicalism to be true about the mind. The nature of subjectivity has been used to argue against physicalism, mainly by Frank Jackson and his knowledge argument; the thought experiment about Mary the scientist (Jackson, 1982), as well as by Thomas Nagel through his essay “What is it like to be a bat?” (Mortal questions, 1979). To solve the problem posed by the possibility that the sciences may never be able to objectively explain subjectivity; every persons (possibly) unique experience of the what it is likeness of an experience, philosophers like David Chalmers have come up with a kind of property dualism that stays within the metaphilosophical account of naturalism, yet manages to slip out of the tight grip of physicalism about the mind. While not explicitly denying physicalism about the mind, Chalmers and the like separate the properties of consciousness in two. The first group of properties abides by the laws of physics and work presumably in accord with traditional physicalism, causality and determinism. Figuring out exactly how this part of our

consciousness works is what Chalmers call the easy problem. The second part of our consciousness is the hard problem to solve. Because this part of consciousness- the phenomenal consciousness has fundamentally different properties, it does not have to abide by natural laws. Somehow, because of the intrinsic specialness of what it is like to drink coffee etc., these philosophers find it necessary to divide what they call phenomenal consciousness from the rest of everything in the world. Now, Chalmers does not pick a side in the free will debate, but this kind of thinking can be seen as the loophole sorely needed to make sense of freedom in a deterministic universe.

Güven Güzeldere addresses the two distinct intuitions we have in the mind- body problem. The first is the one where it seems unlikely for science to ever be able to address the “what it is likeness” of the smell of coffee etc. The other one is an intuition that it makes sense to keep searching for a scientific understanding of ourselves and the rest of the world. I agree with him when he says “it is the second intuition about commitment to a scientific understanding of the world that is to be honored. And, as such, the first intuition about the gap between the phenomenal and the physical descriptions of the world is to be regarded as a challenge to the project of using explanatory means available and looking for others in innovative research programs to come up with more satisfactory accounts of the phenomenal aspects of our conscious lives.” (Güzeldere, 2002). I would only like to add that even if it turns out that science cannot address phenomenal experience objectively, that would not prove that such experiences have properties that are not all physical, but simply that not all things can be explained objectively, maybe you have to actually be the bat to know what it is like to have a bat life. I do not see how a failure of science to give you the bat experience from a book can show that the bat experience has properties that are not like other physical phenomena.

Paul Churchland claims in his book *The Engine of Reason, the Seat of the Soul*, that phenomenal consciousness might be explainable for the science of the future, and that “..even a brilliant thinker can come to assume that what transcends his imagination, transcends discovery by science.” (Churchland, 1996). One of the parallels he draws is the puzzle of life and how it was solved by the discovery of DNA and “a purely materialist, reductionist account of the nature of life- of self replication, of genetic diversity, of evolution, of protein synthesis, of developmental and metabolic regulation- seemed to many scientists to be all but in hand”. And he suggests that a scientific explanation might be available for consciousness as well. This may of course be the case and I do not oppose the possibility that subjectivity

can be objectified, but (it is not certain that science will accomplish this and) it doesn't seem necessary for the physicalist view of the world.

To sum that up, phenomenal consciousness and subjectivity about experiences tend to come up in discussions concerning free will and determinism. However, if we agree to move only within naturalism and physicalism, then these objections or loopholes don't make sense and can safely be discarded in the debate between compatibilists who hold naturalism and physicalism to be true and hard incompatibilists who do the same.

## **1.8. Compatibilism**

Compatibilism is the default take on the free will problem and as mentioned endorsed by 59,1 % of philosophers who participated in the survey conducted by Bourget and Chalmers (2013). It allows for human freedom and determinism to exist side by side in harmony, often accompanied by moral responsibility as well. In fact, very often moral responsibility seem to be what compatibilists are advocating, more than freedom itself as they tend to be on a quest to find/ define a freedom that is just "free enough" to provide moral responsibility and to justify our practice of holding each other responsible in morally loaded situations.

As many will discover/ interpret from exploring libertarianism, indeterminism isn't all that great of an alternative to determinism when it comes to the possibility for true freedom. It seems difficult to perform free actions without sufficient causes for them and much easier when one has at least one good motivation along with control of making the action happen.

### **1.8.1. Classical compatibilism**

The classical compatibilist definition of free will is "Free will is the unencumbered ability of an agent to do what she wants" (McKenna 2009). This view of freedom could be compatible with determinism but is not held by many people today. It's easy to agree that, to use McKenna's example, a mentally unhealthy person could do what she wanted under hallucinations and that doesn't seem to be actions of free will. For instance, she could believe that she was Jesus and needed to die and resurrect in order to save humanity. This could



make her want to commit suicide, something she would not have wanted had she not been under hallucinations (and psychosis).

Determinism is not a threat to our ability to at least sometimes do what we want. In general, the ability to do or choose otherwise is not something compatibilists endorse. The ability to do otherwise in the garden of forking paths is by many compatibilists seen as a threat to freedom. There are exceptions to this rule. Some compatibilists do believe in a sense that we can do otherwise, though this is given that the situation was different in some sense that made the agent want something different. - As a result they are not really talking about the same kind of ability to do otherwise that the libertarians believe we have and that other compatibilists as well as hard determinists deny.

#### **1.8.1.1. Van Inwagen and the consequence argument**

Peter van Inwagen introduced the term incompatibilism as opposed to compatibilism. Incompatibilists are everyone who considers freedom incompatible with determinism (libertarians, hard determinists and hard incompatibilists). He is himself an incompatibilist (libertarian) philosopher and his challenge to compatibilists have to do with the garden of forking paths understanding of freedom and how that is incompatible with determinism. “It seems to be generally agreed that the concept of free will should be understood in terms of the *power* or *ability* of agents to act otherwise than they in fact do. To deny that men have free will is to assert that what a man *does* do and what he *can* do coincide. And almost all philosophers agree that that agent *could have* refrained from performing that act.” (Van Inwagen, 1974) as we have seen, there are some differences between classical compatibilism and modern compatibilism and most modern compatibilists do not advocate the ability to do otherwise. So Van Inwagens challenge can be seen as primarily posed to classical compatibilists.

Today, the philosophers who believe in this kind of freedom are mainly to be found in the libertarian camp, not with the compatibilists. Van Inwagen has been very influential to the free will debate however, so I will give him and his consequence argument some attention. It is more directed toward the classical compatibilists than the modern ones. One formulation of the consequence argument sounds like this: “If determinism is true, then our acts are the consequences of the laws of nature and events in the remote past. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore the consequences of these things (including our own acts) are not up to us.” (Van Inwagen,

1983). Classical compatibilists must refute the consequence argument and one of their strategies are as Kane points out an objection to the statement that our own acts are not up to us. Kane boils this down to an analysis of the word *can*. “Can” may mean the power or ability to make something happen, hypothetically or potentially or actually. For instance, if agent x is at one of the points in the garden of forking paths where he must either continue to the right or to the left, it seems like he *can* chose either- that he has the potential, the ability to make either choice. He does not have a strange brain condition that makes him ignore one side of his visual field or someone holding a gun to his head forcing him to pick one over the other. Even though the compatibilist recognizes that what the agent end up choosing is what he has to do in that particular situation because of the truth of determinism (when it comes to human actions anyways), he objects to the claim that the agent could not make the other choice because surely he could if his motivation had been different.

The hard incompatibilist should in my opinion appreciate the importance of clarifying the distinction between potentiality and impossibility, especially when it comes to the psychology of people who actually believe in hard incompatibilism- it would be harmful to start believing that because you fail at something once that means that you cannot ever succeed in that particular activity etc. But then again one would have to misunderstand hard incompatibilism pretty thoroughly to begin with to believe that. The fact that I chose to walk to the store instead of taking my bicycle does not mean that I had lost my ability to ride a bicycle, but if determinism is true, then the inner and outer circumstances of my situation meant that in that particular situation I could not have taken my bicycle to the store. The ability to perform an action is not enough for the action to occur. A severely clinically depressed person still has the ability to ride a bike and could probably do so at gunpoint, but without motivation (hopefully for the depressed person a gun would be able to spark a tiny hint of it), it is impossible to perform an action that requires will. (Activities like breathing are obviously not dependent on motivation). According to Kane, compatibilists state that the consequence arguments claim that we cannot now do anything other than what we in fact do, is false because of the potentiality we have of at least doing some things differently (like raising our hand or refraining to do so). The way I see it, it is hypothetically true that Mary can raise her hand at the next moment, but factually false if her motivation is not there for her to raise her hand at the actual moment.

Kane describes both types of compatibilists in his Contemporary Introduction to Free Will.

“Many compatibilists today do concede that the *classical* compatibilist analysis of “could have done otherwise” may be flawed, for the reasons just given or for other reasons. But these same modern compatibilists insist that defenders of the Consequence Argument are begging the question when they assume that *no* compatibilist analysis of “could have done otherwise” could possibly be right, merely because the classical compatibilist analysis is flawed. Perhaps this is so. But then the burden of proof lies with compatibilists to give a better account of “could have done otherwise” than classical compatibilists have offered—or to find some other way to refute the Consequence Argument.” (Kane, 2005, p.31).

### 1.8.2. Modern compatibilism

The modern compatibilist conception of freedom is basically and generally the ability to act in accordance with what we identify as our own values and character traits. However, it is possible to imagine a case where a person acts according to her own values and character traits but where we feel weird about calling the action free. Frankfurt’s account of free will is to have your first order volitions or desires match or get along with your second order volitions. For instance if my want for chocolate is my first order desire but my second order desire is to have a ripped body, then I act freely if I abstain from the chocolate. And if I’m not able to act according to my second order desire or volition then my actions are unfree or compulsory. Kane (2005) brings up B.F. Skinner’s *Walden Two* from 1948 as an example of how one can act according to one’s first and second order desires and wholeheartedly do what one wants and yet it seems like this is not freedom. In *Walden Two* the members of a utopian community live very happy lives doing exactly what they want according to their own values and character traits. They can do whatever they want without causing problems for each other because scientists have behaviorally modified them to have unproblematic wants and wishes (Skinner, 1948).

My impression is that most modern compatibilist theories of freedom do not allow for the ability to do otherwise or for the agent to be the ultimate source of himself. Again, there are some exceptions to this but those would be in the form of the classical compatibilists who defend the ability to do otherwise but not when “all things are equal”. Because the garden of forking paths model of freedom is strictly speaking of choice situations where all things *are* equal, the compatibilist strategy to defend this particular type of freedom fails. The ability to

do otherwise is however not needed for the kind of freedom compatibilists like Dennett consider worth wanting, and neither for moral responsibility according to Frankfurt. It also seems that most modern compatibilists believe that human actions are determined by antecedent events just like other macro objects in nature. In order to fit freedom into a determined universe, modern as well as classical compatibilists have come up with many different concepts and definitions of freedom. These are generally versions of the source model of control, though none claim that we can be the ultimate source of our actions unless the word “ultimate” is seriously compromised. As mentioned, the source model of control is the idea of the agent as producer or source of her actions. It is not necessary for her to have the possibility to act differently as in the garden of forking paths model, her freedom lies in her being the ultimate source of her actions or, if her actions are causally related to something outside of her; the *relevant part* of her action was produced by her. The immediately obvious problem is that all of her actions are causally related to events outside of her if one holds causality to be true. The versions of the source model of control types of freedoms many modern compatibilists endorse center around theories of how one can be enough of the source of ones actions to be meaningfully held responsible for them, not their ultimate source.

#### **1.8.2.1. Modern compatibilists strategies for refuting the consequence argument**

Per haps thanks to Frankfurt, modern compatibilists have to a large degree stopped worrying about the garden of forking paths and instead reflected on what makes an action free in the sense that we can be morally responsible for it.

The source model of control kind of freedom where one is the source of one self would do the trick, but most philosophers, even libertarians, do not believe in the possibility of this version of freedom. Alterations of it however, where the agent is not the ultimate source or creator of himself, but where we (or compatibilists at least) still can say that the agent is somehow “enough” of a source for his actions is the kind of freedom most modern compatibilists defend today.

As mentioned, Frankfurt and what Kane calls hierarchical motivation theories are popular with compatibilists today and generally claim that if higher order desires/ motivations match with lower order wants and wishes then the resulting actions are free, whereas if one has poor self control and fails to study for the final exam in medical school, one will not reach the object of ones higher order desire to be a doctor and the decision to go to the beach instead of

studying is an unfree decision. A higher order desire is a desire to achieve or become something that the agent finds very valuable, like a long-term goal of becoming a pro football player or buying a house. While having one of these higher order desires we may experience wants that can be conflicting with the long term goal, such as a desire to go shopping too often when one should be saving for the house, or spending too much time watching TV instead of practicing football. According to hierarchical motivation theories the more you act according to what is most characteristic of you, or in tune with your values and the person you want to be, the freer those actions are.

While the consequence argument is important and influential in the history of the free will debate, it is my impression that compatibilists spend more time coming up with alternative concepts of freedom than the classical “ability to do otherwise”, then they do worrying about hypothetical and factual meanings of *can* and *could*.

In the following sections I’ll describe a few varieties of compatibilist arguments from Dennett and Frankfurt.

### **1.8.3. Compatibilism according to Daniel Dennett**

Daniel Dennett is a popular advocate of contemporary compatibilism. He describes Nicholas Maxwell’s (1984, Dennett 2003) definition of freedom; “the capacity to achieve what is of value in a range of circumstances.” as “as good as a short definition of freedom could be.” (Dennett, 2003, p. 302). If we go along with such a definition of freedom then I agree it solves the problem of compatibility between freedom and determinism, but the baby is thrown out with the bathwater. I can’t see how that kind of “freedom” can have any value in terms most of us (I think) expect freedom to have; grant us the opportunity to control our choices and give us the kind of moral responsibility we can only have in true choice situations. It is like arguing that “yes we do have unicorns in the garden, only they might really be horses with card board toilet paper tubes glued to their foreheads, but they *are* unicorns!”. Or perhaps more precise: “Unicorns exist and are horses with card board toilet paper tubes glued to their forehead”. Point being, the magic that is associated with unicorns is not associated with horses with ugly head ornaments. “Freedom” is the kind of word that, along with “unicorn”, “magic” and “God” do not refer to something in reality but to made-up concepts. And though I do not believe the concept “troll” refers to creatures existing in reality, I would still be able to spot a fake one walking past me on the street.

There is a competition between hard and soft determinists (compatibilists) going on where both parties make arguments about how the other party really agrees with them, and that the only reason why they think they disagree are small and terminological. Dennett, the ultimate champion for the compatibilists, writes in his *Freedom Evolves* that

The hard determinists among you may find ... that your *considered* view is that whereas free will- as you understand the term- truly doesn't exist, something *rather like* free will does exist, and it's just what the doctor ordered for shoring up your moral convictions, permitting you to make the distinctions you need to make. Such a soft landing for a hard determinist is perhaps only terminologically different from *compatibilism*, the view that I am defending in this book. (Dennett, 2003, pp. 97, 98 (in parenthesis)).

There are two problems with this attempt to sneak hard determinists onto the compatibilist team. The first one is that yes, something rather like the type of free will that Dennett suggests probably does exist. I would call that control instead of free will because I find it a much more appropriate term. The type of freedom that Dennett suggests is (as mentioned) Maxwell's definition: "the capacity to achieve what is of value in a range of circumstances." (Dennett, 2003, p. 302). This type of definition doesn't contradict determinism but it also doesn't put humans in the special moral position Dennett wants us to have. Bacteria can get what is of value to them in their circumstances.

The other problem is his claim that this soft landing is what incompatibilists need for shoring up their moral convictions. The hard determinist/hard incompatibilist doctor ordered a change of moral convictions, a small but in my view crucial change. The point is that the hard incompatibilism I hold and many versions of hard determinism are pursuing something quite different from what the compatibilists like Dennett are; they want to keep the moral convictions that allow us to condemn each other, "we" want to remove them.

Jeremy Randel Koons tries in his paper "Is Hard Determinism a Form of Compatibilism?" to show that the answer to the question in the title is yes; hard determinism is just another version of compatibilism and because not many philosophers value libertarianism very highly any more, compatibilism is the last standing contestant aka the winner of the free will debate.

His argument is based on the fact that humans have certain kinds of actions they find praise- or blameworthy and (somehow) he draws from that that humans have some actions

that are free. His “logic” is that because the actions that are considered morally charged are largely the same actions that are considered free, they must also be free and we can all lay our pens down and finally stop debating this topic. He feels certain that hard determinists, given their way with society will continue to punish wrongdoers as the compatibilists and libertarians have. This may or may not be the case and in part two I will look into Strawsons reactive attitudes etc. My point for now is that treating people as blameworthy doesn't show anything of them deserving that or not. If sanctioning is a social way of regulating behavior in a group, then it can be true that some actions are sanctionable even though they are not free. Punishment for the sake of revenge however will not be very meaningful.

### **1.8.3.1. Dennett's Evitability**

Physical laws often predict causation. Laws about the relation between events involve causality. For example; event p causes event q. many natural laws seem to involve causality, this is especially the case for laws in biology, e.g. the function of the heart is to cause blood to be pumped and spread in the body etc.

If we can assume causation among the biological events of human lives, then we can also assume determinism. Compatibilists like Dennett already do that so this is not where our quarrel lies.

Inevitability seems to follow logically from determinism. Daniel Dennett thinks it doesn't and I will try to make a counter argument. His argument goes like this; “In some deterministic worlds there are avoiders avoiding harms. Therefore in some deterministic worlds some things are avoided. Whatever is avoided is avoidable or evitable. Therefore in some deterministic worlds not everything is inevitable. Therefore determinism does not imply inevitability.” (Dennett, 2003, p. 56).

My answer to Dennett's argument is that first of all; evolution has made sure that all living animals today do avoid harm as far as what is possible, if they didn't they would not survive the natural selection. Secondly, given that avoiding harm is possible: it is on the contrary rather (almost) impossible for the organism not to avoid it with its best effort. (Whether they succeed in avoiding harm is of course dependent not just on their own abilities but of the situation, abilities of the harmful thing etc.) There are a few exceptions to this, for instance a suicidal human being, a dehydrated antelope running the risk of crocodile attack to drink water etc. (though you could still make the argument that they are both avoiding harm- for the suicidal person life might seem harmful and for the antelope death is certain without water. Or when it comes to the suicidal person one could make the argument that in order to

commit suicide there must be something so wrong with him or her that they are not fit for survival- and that leaves one less exception to the above “rule”).

The point is if determinism is true, then everything that happens has necessary causes for happening and could not be avoided in the particular situation it occurred in. The reason consists of prior events that had direct causal effect on the happening in question. There is no room for Dennett’s evitability; whichever event has the antecedent events required to happen *will* happen. Again I’d like to point out that this does imply that fatalism is true, nor that perfect predictability will be possible in the future for the reasons given previously.

#### 1.8.4. PAP and Frankfurt

In his article “Compatibilism”, McKenna devotes some space to Frankfurt’s attack on “PAP: A person is morally responsible for what she does do only if she can do otherwise.” (McKenna 2009). This includes an example of a situation in which a person could not do otherwise, but seems (in some circumstances) to yet be morally responsible for his actions. The example according to McKenna goes like this:

Jones has resolved to shoot Smith. Black has learned of Jones's plan and wants Jones to shoot Smith. But Black would prefer that Jones shoot Smith on his own. However, concerned that Jones might waver in his resolve to shoot Smith, Black secretly arranges things so that, if Jones should show any sign at all that he will not shoot Smith (something Black has the resources to detect), Black will be able to manipulate Jones in such a way that Jones will shoot Smith. As things transpire, Jones follows through with his plans and shoots Smith for his own reasons. No one else in any way threatened or coerced Jones, offered Jones a bribe, or even suggested that he shoot Smith. Jones shot Smith under his own steam. Black never intervened. (McKenna 2009)

So Jones shot Smith without the need for Black’s interference even though if he had decided not to do so, he would still have done it under influence of Black. This seems to be a successful counter example to PAP. According to McKenna, “If Frankfurt's argument against PAP is correct, the free will debate has been systematically miscast through much of the history of philosophy. *If* determinism threatens free will and moral responsibility, it is *not* because it is incompatible with the ability to do otherwise. Even if determinism *is*



incompatible with a sort of freedom involving the ability to do otherwise, *it is not the kind of freedom required for moral responsibility.*” (McKenna 2009).

It seems to me that Frankfurt’s argument against PAP is a good argument to show that moral responsibility is not necessarily eliminated by the absence of the ability to do otherwise. It does not however prove that Jones was morally responsible for shooting Smith. If that was what he meant to prove than Frankfurt may not be critical enough toward the source model of control. It is not a given that we have moral responsibility simply because we may want to kill someone of our own accord. Jones acted on his own unencumbered because Black did not manipulate him to shoot Smith. According to causality, causes do not pop up from nowhere inside of agents; every event is part of a causal chain whether it is described as a human action or as photosynthesis. As an alternative to the ability to do otherwise, the source model of control may not be any better, or more compatible with determinism if interpreted in the sense that the agent has to be the ultimate source of his or her actions. With that said, I don’t want to accuse Frankfurt of relying on the source model of control in any simple way as he his volition theory is a complex type of source model of control that does not as far as I’m aware of claim to be ultimate. I get back to his theory of freedom. As for the moral irrelevance of an ability to do otherwise McKenna might be right, but there is no need to go there. The kind of freedom relevant for moral responsibility is even more farfetched than the ability to do otherwise. I will get back to this in part two of my thesis.

### **1.8.5. Manipulation arguments**

As mentioned, modern compatibilists seem to mostly focus on the source model of control, rather than the garden of forking paths model of control. If the source of an action can be said to be the agent, in some crucial way, then the agent owns that action and has acted freely. The “crucial” way might be what Frankfurt calls the second order when it comes to desires and wishes. The problem of the alterations of the source argument which many compatibilists base their beliefs of freedom on is the fact that if determinism is true, then no causal chains of events really starts within agents. While compatibilists are aware of this and try to avoid the problem by saying that an action only has to be caused by an agent in a “crucial” way, not ultimately, this dilemma still invites a series of arguments that have been called “manipulation arguments”.

It is intuitive to most people that if someone have been thoroughly manipulated by someone else, they should not be held morally responsible for what they have done under this manipulation, and that they have not acted of their own free will while performing acts in the interest of the manipulator. These arguments can involve hypothetical stories of brain washing, mass socialization like in *Walden Two*, or more intricately; neuroscientists who kidnap people and place chips in their brains (or other sci-fi style devises to impose brain control), and remote control people to do things they would not normally do on their own accord. These scenarios are then compared to real situations in life where people have done something morally condemnable, and which most compatibilists would claim the agent was morally responsible for because the action was done consensually with their values and / or character traits. One might say for instance, that a person growing up in a religious community and living his life according to the values he learned while growing up is fully capable of acting freely if he has no pathologies etc. A person growing up in a strict, controlling and brainwashing sect however, would not necessarily have the option to act freely of his or her own accord because this would be so pressured and controlled by the community that the compatibilist would not sit comfortably holding this person morally responsible for something he or she did that was coherent with the crazy beliefs of the sect.

The problem for compatibilists is that if you do admit that no one are the ultimate causes for their actions, but insist that some can be held morally responsible for their actions while others must intuitively be spared this harsh judgment, how do you really separate the two? How is being manipulated different from being a part of natures system of cause and effect when it comes to moral responsibility?

Michael McKenna sums the manipulation argument up like this: “(1) Any agent so manipulated into *X*-ing is not free or morally responsible, despite satisfying the pertinent compatibilist-friendly conditions. (2) Any agent who is determined to *X* is no different in any relevant respect from an agent manipulated into *X*-ing in the manner specified. Therefore, (3) any agent who is determined to *X* (or, generalizing, perform any other action) does not *X* freely and is not morally responsible for *X*-ing.” (McKenna, 2009).

#### **1.8.5.1. Compatibilist replies to manipulation arguments**

As a reply to manipulation arguments, compatibilists may object to the premise number one and claim that an agent can in fact be free and morally responsible while under manipulation. Another way to deal with the problem is to object to the second premise and claim that yes

there is a difference between being determined and being manipulated. In my view this is also the best strategy for dealing with the manipulation argument and I think most people would certainly prefer to be determined by the past and natural laws etc. than by some maniac individual, sect or “big brother” government. While anyone can appreciate the difference, explaining how the moral responsibility fits in is still more difficult. The reason why this difference between determinism and manipulation is relevant for moral responsibility is according to compatibilists because freedom is at least compatible with determinism if not dependent on it. Manipulation however is not compatible with freedom and there you have it.

These reply strategies do not seem to work, the first for what I daresay are obvious reasons and the second because still seem to lack a good account of how exactly one can be “free enough” to be morally responsible while performing determined acts.

### **1.9. Hard determinism**

The hard determinist believes, just like the compatibilist, that human actions are determined by antecedent events outside of our control. Unlike the compatibilists, they do not accept the possibility for freedom in a world where human agents cannot “do otherwise” and quite certainly haven’t made themselves.

Firstly, I explain what I take hard determinism to be, based on my knowledge of the tradition. Laws of nature seem to confirm determinism, at least at the macro physical level.

Hard determinists consider the relationship between cause and effect as necessary in the sense that if x causes y on Tuesday, then all things equal, x will necessarily cause y on Wednesday as well. This is a refutation of the Humean account of causality and newer theories of causality as series of events that tend to cause other events- but do not do so necessarily (Mumford and Anjum, etc.). With this view of causality it is to the hard determinist logical to assume that humans cannot have freedom when everything we do and are is strictly dependent on events of the past that are clearly outside of our control.

The reason why hard determinists are not compatibilists has to do with their understanding of what freedom is. Hard determinism is one of the two anti- freedom positions in the free will debate and is often portrayed as a pessimistic and hopeless view on humanity and that believing in it, regardless of it being true or false, contributes negatively to humanity. There is a mixture of reasons for thinking this; (perhaps) most importantly the western protestant

cultural heritage that glorifies hard work and self made men, second chances etc. It can also be caused by understanding it as fatalism, the view that people cannot help themselves from what's going to happen because everything is outside of their control. Que sera sera.

It is also a common worry among philosophers that without a belief in free will, we will all run amok and rape, kill, steal and ruin civilization. It is true that according to hard determinism humans do not ultimately control things because that would require freedom- which hard determinists don't believe we can have in a deterministic world. However, the lack of freedom does not mean that one cannot help oneself to a better life, merely that in order to do so, one must have some kind of cause to change his or her course in life. Human actions are dependent on motivation, which in turn is dependent on a myriad of feelings, desires, beliefs etc. So the hopelessness of hard determinism is not really accounted for and in part two I will come back to why the fear of ethical apocalypse is not justified either.

Coming into this, hard determinism was the position I favored until about half way through the writing process. I discovered that my motivation was to argue against "ultimate moral responsibility" and the practice we still have to condemn each other in cases where an individual has committed a particularly harmful crime or deed. I found myself in numerous conversations trying to defend the thesis of determinism - which I believe in if understood as human actions being determined, but which is not important for refuting the kind of freedom that can possibly provide us with the kind of moral responsibility I initially do not think we have. I have interpreted that science claims that human actions are determined, but not the entire universe- that there is some indeterminism on the quantum physical level etc. If this interpretation is correct then one can hold the hard determinist position in the free will debate if one wants to, but it is not necessary to do so for what I wish to claim. It also seems insignificant for the possibility for the kind of freedom I'm talking about; indeterminism definitely does not seem to provide us with any more control than determinism. Here I'm prone to agree with the compatibilists who claim that indeterminism would make us less free than determinism because it would subject us to chance in addition to the causal effect of antecedent events. It is not clear to me whether or not hard determinists would be libertarians if the thesis of determinism were thoroughly refuted. The way I see it, the libertarians are the ones who have the best notion of freedom, but it is regrettably not available to us neither with nor without determinism.

### 1.10. Hard incompatibilism

The position of hard incompatibilism is the simplest position in the free will debate. It claims neither that determinism is true nor that freedom depends on it, and neither that indeterminism is true. It is a small position, 12,2 % of philosophers asked in Chalmers and Burges survey reported to believe in “no free will” and I presume most of these are likely to be hard determinists.

Basically, hard incompatibilists figure that by definition, freedom is something that collides with human nature, regardless of determinism being true or not. The term hard incompatibilism can sometimes be used to describe hard determinism and sometimes to describe someone who denies freedom and moral responsibility. My version will be of the latter and I separate this view from hard determinism mainly for the reason already mentioned; the fact that the kind of hard incompatibilism I endorse does not need to worry so much about determinism, which is the foundation of the hard determinists beliefs.

Derk Pereboom introduced the term hard incompatibilism for claiming that freedom and determinism are incompatible and that true free will does not exist. As opposed to hard determinists, hard incompatibilists do not depend these claims on the truth of determinism. The term hard incompatibilism can sometimes be used to describe “modern hard determinists”; philosophers who have the hard determinists opinions about freedom and moral responsibility, but less of an emphasis on determinism as it has technically been refuted by science (technically because science has not refuted the theory that human actions are determined, merely the claim that all events in the entire universe are determined, quantum mechanics etc.). In my opinion, if one is talking about modern hard determinists it is better to use the term hard incompatibilism because of the fact that determinism is less in focus with modern freedom deniers and therefor it is disadvantageous to have the word determinism in the name of the position.

Advocating this position requires an argument for why we should choose the hard incompatibilist definition of freedom and not one of the various compatibilist definitions. So the definition goes something like this; “to truly have freedom, a human agent must have the ability to do otherwise in a “choice situation” or at least be the source for his or her choice”. In Kane’s words, one must have freedom of self-determination and /or freedom of self-formation (see 2.1.1.3). If we believe in metaphysical naturalism and physicalism, and that there are causal chains of events leading up to every persons life and every event in it then it seems clear that no human can be the source of his or her choices or actions and thus cannot

possibly have freedom. Now, why should we want to accept these extreme demands on the concept of freedom? Why not gain freedom by giving it some less impossible content?

The various compatibilist definitions of freedom have in my eyes failed to show how they allow for ultimate moral responsibility. If a freedom worth wanting is a freedom that provides us with ultimate moral responsibility, at least for some actions, then settling for a definition of freedom that does not satisfy in providing ultimate moral responsibility is to trick oneself.

Now, having the opinion that ultimate moral responsibility can only occur when libertarian freedom is present might be a kind of personal preference and it may or may not be possible for compatibilists and incompatibilists to come up with a solution everyone are happy with. About the dissatisfaction incompatibilists might experience about compatibilist definitions of freedom Kane has this to say:

“If you are still not satisfied with the above account of freedom of will, then it is no doubt because you are thinking of free will in some further sense than simply the ability to choose or decide *as* you will without constraint. You must be thinking of freedom of will in something like the ‘deeper’ sense of free will of chapter 1—as a kind of *ultimate* control over what you will or want in the first place: A control incompatible with your will’s being determined by any events in the past over which you did not have control. Now we compatibilists obviously can’t capture *that* deeper sense of freedom of will, no matter what we do, because it is incompatible with determinism. But, as compatibilists, we believe that any so-called deeper freedom of the will—or any kind of free will that requires indeterminism—is incoherent anyway. No one *could* have a freedom of will of such a deeper kind.” (Kane, 2005, pp. 15- 16).

This is the core of the issue between compatibilists and incompatibilists and it can seem that different intuitions about freedom may stand in the way of ever agreeing. The compatibilists will rather have their limited freedom and unaccounted for moral responsibility whilst the incompatibilists who are not libertarians agree with Kane that freedom of a “deeper kind” is something we cannot have, and that freedom has to be “deeper” in order to be freedom and not just control- which we don’t have the same problems of seeing the limits of.

### **1.10.1. The hard incompatibilism of Pereboom**

Derk Pereboom is one of the most prominent philosophers who endorse hard

incompatibilism. In his book *Living Without Free Will* he describes his own account of hard incompatibilism as consisting of two thesis's. The first is the thesis that human actions are either determined by factors outside of their control or by truly random events that have no cause or a mix of the two. The second thesis is that the "Causal History Principle's" way of defining hard incompatibilism is true. (Pereboom, 2001: 127) The Causal History Principle states "An action is free in the sense required for moral responsibility only if it is not an alien-deterministic event, nor a truly random event, nor a partially random event".

Pereboom considers agent causality metaphysically and conceptually possible, but thinks there are good reasons to assume that we do not in fact have this kind of freedom.

### **1.11. Conclusion of part one**

Defending hard incompatibilism from opposing arguments coming from the other positions in the free will debate has been done throughout part one while criticizing the alternative positions. I have argued that the term freedom should in this context be the libertarian agent causal understanding of free will and that ultimate moral responsibility and freedom should be considered connected.

The version of hard incompatibilism I have come to hold in the process of writing this thesis is one that differs a bit from the view of Pereboom. The distinction between what I've called *ultimate moral responsibility* and *practical moral responsibility* is to my knowledge not found with Pereboom. But like hard determinists he does separate morality from moral responsibility and he claims (without much protest from the other contestants in the free will debate) that hard incompatibilism is not incoherent with morality, only with moral responsibility.

The difference between the Pereboom's view and mine is not massive, but significant. I agree with him and the Causal History Principle that we cannot have the kind of freedom required for moral responsibility only if we are talking about the ultimate kind. In the form of morality and the ability to be responsible, practical moral responsibility is perfectly possible even if our actions are alien- deterministic-, truly- or partially random events. I will get back to this in part two.

## 2. Part Two: moral implications of hard incompatibilism

### 2.1. Introduction

In this part of my thesis I'll deal with possible/thinkable consequences of the conclusion I drew in part one. My assumption is that in the hypothetical scenario where such a conclusion were to be publicly accepted, the impact of it in society would be small but positive given that people correctly understood it (this would of course be pragmatically problematic).

The main statement of this part of the thesis is that if we have a naturalistic understanding of what morality is, and with the definition of freedom given in part one, then it follows logically that we do not have the kind of moral responsibility (ultimate) that presupposes freedom and can justify condemnation. By "naturalistic understanding" I mean to view morality not as something that sets us apart from other animals (not as something God given or transcendent from something outside of nature), but rather one of the characteristics of the kind of animal we are. I claim that morality is something we have independent of freedom, and that we are practically morally responsible when we are able to be so.

Two important questions in this part are 1) if we do not have ultimate moral responsibility, why then is it that we individually have such strong intuitions about being morally responsible? And 2) how can we make amends for the fact that people at the end of the day cannot be held morally responsible for their actions?

I will try to answer the first question through an investigation of the concepts of morality and responsibility. The answer will basically be that we have those intuitions because we are able to be practically morally responsible a lot of the time.

A tentative reply to the second question will be made up of reflections on the difference (if there is one) between on the one hand, holding someone responsible according to what P.F. Strawson says about reactive attitudes etc., and on the other hand, what it means to be held ultimately morally responsible.

I will also address the widespread worry of what will happen if people were to believe the conclusion of my thesis, and I will base myself on Strawson and Nichols for this.

Then I look into the philosophical problem of moral luck, brought to popular attention by Nagel. Does hard incompatibilism contribute to the unfairness of our judgments or might it be a better alternative, at least with respect to the problem of moral luck than say compatibilism?



The next section is devoted to what Clarke calls “heaven and hell” responsibility; different philosophers have used the idea of hell as an illustration of what ultimate responsibility means in its most “ultimate” sense, at “the end of the day” what can we be held responsible for? The weakness of this illustration is that modern humans should not worry about or believe in hell anymore and the strength is how it really puts things on the edge and illustrates a difference between practical moral responsibility and ultimate moral responsibility.

Toward the conclusion of this part, I’ll bring up some reflections on criminal care and how hard incompatibilism may be complementary to the development of the justice system in Norway and perhaps even something, which could further encourage it. If we believe that humans aren’t ultimately responsible for who they are it might be easier to look for explanations of how some individuals end up on the outside of society.

## **2.2. Morality- naturalistically understood**

Descriptively speaking, morality is the code of conduct observed in groups or societies of humans and arguably to an extent in some other animals like (but not restricted to) primates. When it comes to the human species, I believe there are good reasons to hold that morality is universal more than relative. This has to do with my naturalistic understanding of what morality is, and in the following I’ll try to explain it.

A naturalistic understanding of morality builds on the perspective of humans as social and intelligent pack/ group animals. Central to speculations about why we have developed morality is our immature and dependent offspring and our need for cooperation. This is a reductionist portrayal of morality and only semi scientific speculation as it has not yet been finally decided exactly what brings about our moral behavior and where “moral activity” is located in the brain etc. The point is to try and understand morality as something that makes sense based on what we are- naturally speaking.

So, while being careful about exactly how we are wired with respect to morality, I am speculating that morality is the sum of abilities like sympathy, empathy and the capacity to abstract thought (in this context for the formation of principles, norm and rules). And not as a byproduct of these abilities, but to serve a much needed function; making survival easier. This understanding of morality runs in the opposite direction of cultural relativism. And even though there are some seemingly major differences in the customs of different cultures, these are really very tiny compared to how alike morality displays itself across cultures and the

differences can often be explained by factors in the environment (like the persuasion/expectation of elders to commit assisted suicide among certain Inuit groups or the accept of infanticide in cultures who regularly experiences hunger).

Empathy and sympathy is traditionally understood as providing the emotional part of our moral behavior. These emotions can probably be further reduced to oxytocin, dopamine and or other chemicals in our brains developed to ensure the survival of babies and the cooperation we have depended on throughout the human evolution. I contend that morality is so similar across different cultures because it has the same springboard: evolution. It is generally forbidden to kill, steal, deceive etc. and generally obligatory to take care of the children one has acquired as well as sick and elderly members of society and to share, reciprocate, help each other etc. A person who follows these “rules” is generally viewed a good person in most or all cultures I daresay. People obviously don`t always follow these rules, but I think it is safe to say that these are general human values that our moral laws encourage and that are only broken under special circumstances.

### **2.2.1. Responsibility**

To be responsible generally means to be accountable for some event under one`s control- to be answerable (able to respond). Moreover, a responsible person is generally viewed as someone who delivers what is expected of him, honors his word and obligations and acts somewhat reliably.

Mark Bernstein states that “Stripped of its legal trapping, we have what amounts to a framework for understanding what philosophers mean when they speak of “really”, “truly” or “ultimately” being responsible for one`s will and the actions that emanate from it. To have free will, in this uncompromised, full blooded way, is to have the power to be the ultimate or final creator of her own intentional mental life” (Bernstein in Wettstein, Fisher, 2005, p.2). He claims that without it we cannot be morally responsible. “To my mind...we can delete not only the scarce quotes around the adjectives that describe the responsibility that concerns us but also omitting adjectives altogether. “True” (“ultimate”, “real”) responsibility is nothing more than true (ultimate, real) responsibility, which in turn, is nothing other than responsibility. There is no error in attaching the modifying clauses or in retaining the scare quotes as long as we are not mislead into believing that we are referring to some arcane segment of our ontology when we do so”. (Bernstein in Wettstein, Fisher, 2005, p.2)

In accord with naturalism and the themes of this thesis, I contend that responsibility is what we may have in situations where we do have a great deal of control- the kind of control Dennett and many compatibilists would call freedom. Besides from this, I think it makes sense to say we have responsibility when we agree to play some kind of game that requires responsible actions from the gamers to work. When getting our license to drive we agree that we have to take the consequences of drinking and driving even though we cannot choose to be the kind of person who ends up doing so. I don't think this is a problem for the hard incompatibilist. It should also not be a problem for hard incompatibilism when some of us end up in jail for a long time for having violated that responsibility. However, when naturalist- hard incompatibilists are talking about moral responsibility I disagree with Bernstein about his protest to applying adjectives like "ultimate" and "true" to moral responsibility. As mentioned, I believe it makes sense to separate between the practical moral responsibility and ultimate moral responsibility.

What I have called practical moral responsibility is what we have in situations where we either agree to take responsibility (as with getting a drivers license), are forced to do so because society depends on it (like paying taxes), or are able to do so in a natural way; like return something we have stolen because we feel guilty about it. What I have called ultimate moral responsibility is something that requires the libertarian kind of freedom I have concluded that we as humans cannot have because of its magical nature.

As Bernstein said, there is no problem with separating responsibility from "ultimate" responsibility as long as we don't think they are ontologically different. It makes sense to call a teenager who gets good grades and is a reliable friend responsible, and one who skips school and regularly ditches her friends and cannot be counted on irresponsible, even though if determinism is true, the responsible teen is no more responsible for being responsible than the irresponsible teen is for not being responsible. I will however claim that there is an ontological difference between the two moral responsibilities. As I have explained, it seems to me that ultimate moral responsibility requires free will which, I'm in this thesis arguing that we do not have. The other kind; the practical moral responsibility does not require free will and if hard incompatibilism is true that does not mean that we do not have moral responsibility of this kind. A world in which we have free will and ultimate moral responsibility is ontologically different from a world in which we do not.

With this understanding of moral responsibility, the practical aspects of responsible behavior stand strong and are not threatened by hard incompatibilism and the absence of freedom. The absence of ultimate moral responsibility and freedom really only hurts the

beliefs of typical religious people who expect people to pay for their sins on judgment day etc.

A common protest against denying that we have free will is based on the concern about what will happen to our moral behavior if we start believing that we do not have the kind of free will most of us today assume we have. As I have argued, freedom and moral responsibility should be seen as two sides of the same coin and this is also in accord with how most of us are brought up believing. We have to be able to blame a person for his or her actions in order to try and make him or her take responsibility for them. It is superfluous to point out why this is a good thing; kids need to learn that they are powerful agents who can do a lot of damage if they don't pay attention to or think through their actions.

As I have mentioned earlier, the denial of freedom and the concept that word represents in my understanding is not the denial of control, agency and the possibility to change one's own life according to what is important for each individual. However it does accord with a denial of the belief that we can have ultimate moral responsibility for our actions.

Many compatibilists tend to think that because of the way we interact with each other, because of our morally reactive attitudes and age-old social system of holding each other responsible for our actions, the denial of free will must either be false or if true; ignored. If it turned out to be true that humans don't really have free will, then that would be a destructive discovery best hushed down. Peter F. Strawson is well known for his essay "Freedom and Resentment" where he carefully investigates what he calls our reactive attitudes in moral situations. In the following, I'll try to present his claims in that essay, along with my interpretation of his main points as compatible with the hard incompatibilism I'm advocating in this thesis. After that I will use the help of Shaun Nichols and his paper "After incompatibilism: A naturalistic defense of the reactive attitudes" to try and strengthen my point.

### **2.3. Strawson and reactive attitudes**

While identifying himself as the kind of philosopher who doesn't know what the thesis of determinism is, Strawson still offers a helping hand to those of us who think we know enough about it to assume it is true or at least likely, in the question of what that should imply for human interaction in situations most of us agree are morally charged in some sense.

When we interact with each other we have reactive attitudes of gratitude, resentment, forgiveness etc. These are reactions to how we are treated by each other in different situations

and based largely on the intention we think the person who does something to or for us is; ill willed or with our interest as motivation; good willed. We can also be treated with indifference.

The reactive attitudes we meet these intentioned acts with are as Strawson says natural. Our considerations of others who do something bad to someone else, causes injury or indifference are divided into two groups. Situations that can inspire statements like “he was pushed”, “they left him no alternative”, “it was the only way” etc. invite us, according to Strawson, to spare the agent of our ordinary reactive attitudes such as resentment. I think he is right about that, at least that it is rational of us to do so.

Strawson again divides the second group of considerations into two subgroups. The first concerns situations where we can say things like “he wasn’t himself”, “he has been under a great deal of strain recently” etc. In these cases we consider the situation so abnormal that we do not judge the person based on whatever he did in the abnormal situation.

The other subgroup of considerations however occur in situations where declarations such as “he is only a child”, “his mind has been systematically perverted” etc. apply. In these cases we consider the agent psychologically abnormal and view him objectively as someone who needs treatment, training or other measures that aren’t intended for sanctioning. We see the agent under a different light than we normally see agents under and “all our reactive attitudes tend to be profoundly modified” (Strawson, 1962, section 4).

I agree with Strawson in this and I want to take it even further. When we know enough of the causes of extreme behavior to understand how it came about, (when we look at people under the objectifying light to figure out how they work) condemning them in addition to the understanding is superfluous.

In what Strawson writes lays the question of what our idea of freedom means for our reactive attitudes. Do we consider them dependent on free will? If Strawson somewhat correctly describes our reactive attitudes, it seems that we at least divide between actions people seemingly have been in control over to some extent, and the actions it seems people have not been in control over. And often, we need more information than our own observation to decide how much control we think a person has had over a certain action. Given that I’m in this thesis assuming that the commonsense intuition of freedom is that we have a kind of agent causal free will, I proceed to suppose that our reactive attitudes would not change much if we did not grant this control to agent causal free will, but rather just easier explainable control and self control.

So to Strawson's question, "What effect would, or should, the acceptance of the truth of a general thesis of determinism have upon these reactive attitudes? More specifically, would, or should, the acceptance of the truth of the thesis lead to the decay or repudiation of all such attitudes? Would, or should, it mean the end to gratitude, resentment, and forgiveness; of all reciprocated adult loves; of all the essentially personal antagonisms?" (Strawson, 1962 p.10). -My reply is of course no to all of it. As he points out, those reactions are natural for humans to have and it is practically inconceivable that we would stop having them based on a theoretical conviction of the truth of the determinism thesis. I assume that reactive attitudes are beneficial for us and that this is why we have them (evolutionary speaking). I can think of two ways they are good to have. One: they signal "don't mess with me" and is a part of the individual's assessment in the group (being a social animal etc.). And two: it teaches, or at least tips off, the wrongdoer to change his behavior. Of course this teaching function doesn't work if the wrongdoer is beyond a place where one feels the impacts of social sanctions. "Pull yourself together!" makes sense to say to someone whom we consider able to do so in the future, but seems a bit absurd to say to a notorious rapist.

Strawson goes on to deepen his question: "Could, or should, the acceptance of the determinist thesis lead us always to look on everyone in this [objective] way? For this is the only condition worth considering under which the acceptance of determinism could lead to the decay or repudiation of participants' reactive attitudes." (Strawson, 1962, p.8, section 4).

Strawson does not think we will change our reactive attitudes given that we start believing that determinism is true, but he doesn't think it is self-contradictory to suppose that that might happen. I find it self-contradictory because of the several times he has claimed reactive attitudes to be natural and that "these practices, and their reception, the reactions to them, really are expressions of our moral attitudes and not merely devices we calculatingly employ for regulative purposes. Our practices do not merely exploit our natures, they express them." (Strawson, 1962: last page, section 6). The thesis of determinism cannot be true and conflicting with human nature. That is logical even if we do not know what the thesis of determinism is.

Strawson also talks about our self-reactive attitudes that are based on demands we have on ourselves for others and that is associated with feeling obliged, guilty, remorseful, responsible etc. These are also examples of reactive attitudes that tell us something about what we expect of others and ourselves in morally charged situations. What I think Strawson wants to say is that the fact that there are reactive attitudes in morally charged situations tells us all we really need to know about the nature of morality and that if the thesis of

determinism is true then it could not compromise the way we build our relationships and deal with each other in our communities.

I agree with him in most of the things he is saying. While I think he is right about it being rational, it seems to me that in practice, it is often not the case that we treat people who fit into the second subgroup of the second kind of considerations (the ones who's minds have been "systematically perverted", "are only children" etc.) in an objective way with profoundly modified reactive attitudes. If we did, I imagine that society would be better and more understanding.

Secondly, if the thesis of determinism is true then I think it should strengthen the rationale of separating the unreliable persons from the rest in consideration and judgment, and that it would not change the way we otherwise interact and have expectations and reactive attitudes toward each other, and neither that it would be a logical consequence of any sort for the truth of the thesis of determinism to provide such a revolution of inter human moral considerations. The fear many philosophers have of the possibility that the thesis of determinism might be true is groundless. According to determinism, all behavior is determined. If an agent does something very anti social and destructive, an action that some people would call evil and others deranged is also determined, but by causes that are not normal for (psychologically and physically) healthy human beings. The causes of such actions differ significantly from the causes that normally determine our actions. As clarified in part one, I'm not denying that our will many times determine our actions but it is assumed here that it is not a free will.

Now, while I do agree with Strawson that our reactive attitudes are very modifiable according to what information we have about the persons who hurt other people (or animals for that matter), I think there will always be a huge difference in actions that harm us or ours personally, and actions we here about in media etc. Surely there are some admirable personalities who forgive their rapist after hearing about his traumatic upbringing, and that's nice, but I think the most natural (-natural in the sense of raw and unlearned) reactive attitude is something more along the lines of hate. Fortunately, the practice of taking an eye for an eye is no longer an option in modern societies, with or without a belief in free will and the personal hate felt by victims of crimes is therefore not relevant.

Sam Harris compares ax murderers to crocodiles and point out that if we stop believing in free will, victims may start to view the people who have done horrible things to them the way some victims of animal attack today may view the crocodile or other wild animal which attacked them; with a kind of awe and excitement. He imagines that the belief in determinism

would allow us to view ax murderers with objectivity and curiosity. (Harris, 2013, [http://www.youtube.com/watch?v=\\_FanhvXO9Pk](http://www.youtube.com/watch?v=_FanhvXO9Pk) )

Shaun Nichols gives in his paper “After incompatibilism: a naturalistic approach to reactive attitudes” a description of why the reactive attitudes that Strawson put the headlights on are perfectly consistent with hard incompatibilism. Nichols divides hard incompatibilists/determinists into two groups: the revolutionaries and the anti revolutionaries. The revolutionaries, as the term warns, are the people who think that the conclusion that humans do not have free will and true moral responsibility should lead to a massive upheaval of how we deal with each other on all levels of society. For instance: it is ultimately unfair to incarcerate a criminal who could not have acted differently than he did. It is also meaningless to reciprocate with anger toward someone who has betrayed you as a friend because he too could not have acted differently. We might feel sad about things not being the way we would like them to be but blaming people for things they do is as silly as blaming clouds for raining. The anti revolutionaries consider the revolutionaries as having heavily misunderstood. They view reactive attitudes as natural and part of the causal system that determines how we act. Thus are reactive attitudes perfectly consistent with hard determinism or hard incompatibilism. Nichols places himself among the anti revolutionaries. “Some maintain that this would usher in a badly needed revolution in our practices. Others worry that the recognition of determinism would lead to catastrophe. I have a more humdrum guess – if people come to accept determinism, things will remain pretty much the same.”(Nichols, 2007, p. 3).

#### **2.4. Revision instead of revolution**

I agree with Nichols on most of his claims. That is to say that if we all suddenly started to believe what Nichols and myself believe right now, not much would change. However, recent history has shown that as the sciences provide more and more detailed explanations for human actions, our reactive attitudes modify (just like Strawson claims they do) and the way society deals with criminality and problem behavior in youths has gone through something that can be called a revolution. Corporal punishment in justice systems is pretty much only practiced by a few Islamic countries that uphold sharia law, the death penalty is absent in most developed countries and in Norway the term *criminal care* is perfectly internalized. Violent or otherwise seriously criminal youths are generally taken care of by child protection institutions instead of thrown in jail and kids who disturb in the classroom are given



individual guiding and extra learning resources instead of being beaten with a cane. Nowadays, there is agreement across the scale of the positions in the free will debate that there is a significant difference between children and adults when it comes to the ascribing of responsibility. While some will want to excuse adults, everyone agree that responsibility should not be ascribed to children. I'm mentioning children because there has been a process of "humanizing" treatment of all "wrong-doers" or "troublemakers", and I suppose that has to do with increased understanding of human behavior.

It can fairly be assumed that the sciences will only continue to produce better and better explanations for human behavior in different situations in the future, and I am guessing that this will continue the trend of humanizing punishment/ substituting it for treatment and rehabilitation whenever possible. This will hopefully happen regardless of the free will problem being solved or not, but hard incompatibilism is in my view the position of the free will debate that matches or compliments this development the most. –Hard determinism could of course do the same job if it didn't have the weakness of insisting that every aspect of the entire universe need to be determined in order for the thesis to be true.

## **2.5. Moral luck versus ultimate equality**

If libertarians or compatibilists are right about morality and we can in fact be ultimately responsible for our actions, then there is an undeniable factor of luck involved in how good or bad we turn out to be. As projects like wartime Nazi Germany have proven, most people are capable of doing the worst of actions. Normal, smart and healthy people can turn their machete on their neighbor or trade and treat humans as livestock (not to mention trading and treating livestock as livestock) if the circumstances are calling for it. A libertarian or compatibilist may concede that he would be one of those very few Germans who saved Jews instead contributing to the carbon monoxide gassing them, or a brave pacifist in Rwanda, but the statistics would be against him, as would psychology and history.

With his essay *Moral Luck* Thomas Nagel coined an important term in the world of ethics. I don't think he is mistaken when he says that before we reflect, it is intuitive for us to think that we should not morally judge people for things they could not help- for events which occurred outside of their control. "Where a significant aspect of what someone does depend on factors beyond his control, yet we continue to treat him in that respect as an object of moral judgment, it can be called moral luck". (Nagel, 1979, p.26).

Nagel makes a distinction between external influences of a close range, factors that lead us to excuse people of their actions quite easily and internal influences that are not as easily noticeable, and less likely to lead us blame or praise people less. Of the external factors he lists “clear absence of control”, “physical force” and “ignorance of the circumstances”. But there are many other, some of which internal influences of a broader range that determine how our intentions succeed-even libertarians can agree to this. So even when our intentions are good, there are always factors beyond our control involved in the situation, and which end up being mixed into the basis of what other people can judge us for. “However jewel-like the good will may be in its own right, there is a morally significant difference between rescuing someone from a burning building and dropping him from a twelfth-story window while trying to rescue him.” (Nagel, 1979, p.25). If Nagel is right, then this adjustment of praise (which is one way to judge) happens when the intention or will was good. That renders little doubt that it also happens when the will was clearly or potentially not good, or even bad. Even if compatibilists are right, there are some parts of the making of our psychology, which is not up to us, and sometimes those parts may be what explain how some of us end up as bad people with ill will toward others.

If determinism is true, (as compatibilists think it is,) then some aspects of our actions are always outside of the scope of our control and thus insisting on holding people ultimately morally responsible will result in the practice of judging people based on luck. Explicitly, Nagel only points to this as a problem and conclusively when he discusses whether or not the problem may have a solution, and he airs the idea that it might be that the incompatibility between the conception of moral luck and our lack of full control of our actions might only appear conflicting.

A solution to the problem of moral luck, which Nagel does not explicitly discuss in his essay, but to me seems hinted about throughout it, is hard incompatibilism about freedom and the downright denial of ultimate moral responsibility. He seems to want to embrace determinism about human actions, though ambivalent about it when he says things like “I believe that in a sense the problem has no solution, because something in the idea of agency is incompatible with actions being events and people being things. But as the external determinants of what someone has done are gradually exposed, in their effect on consequences, character, and choice itself, it becomes gradually clear that actions are events and people things. Eventually nothing remains which can be ascribed to the responsible self, and we are left with nothing but a portion of the larger sequence of events, which can be deplored or celebrated, but not blamed or praised.” (Nagel, 1979, p.37). While expressing

concerns about the self, agency etc., Nagel sounds more like he belongs in the hard determinist or hard incompatibilist camp than any other place in the free will debate and perhaps the only thing stopping him from concluding this was about the problem of moral luck is because of his opposition to reducing consciousness to something physical.

In *The View From Nowhere*, Nagel expresses his view on the free will debate and why he does not participate in it as a defender of one of the positions. “This naturally suggests that the threat is unreal, and that an account of freedom can be given which is compatible with the objective view, and perhaps even with determinism. But I believe this is not the case. All such accounts fail to allay the feeling that, looked at from far enough outside, agents are helpless and not responsible. Compatibilist accounts of freedom tend to be even less plausible than libertarian ones. Nor is it possible simply to dissolve our unanalyzed sense of autonomy and responsibility. It is something we can't get rid of, either in relation to ourselves or in relation to others. We are apparently condemned to want something impossible.” (Nagel, p.112, 1986).

If hard incompatibilism is true and no humans are free or ultimately morally responsible, the element of luck involved in the practice of praising and blaming according to P. F. Strawson won't go away.

The *problem* of moral luck and ultimate responsibility or freedom will however be solved with the denial of the latter two ideas. As many modern compatibilists, hard determinists and my kind of hard incompatibilists agree, Alfred who chooses path a in the garden of forking paths on a Wednesday- all things equal will make the same choice on Thursday. And if our selves- the subjective experience of being me and not you, is physical and we were to somehow exchange a different consciousness with Alfred's, though everything else stayed exactly the same, then the safest bet is that Alfred with a new consciousness inside will make the same choice as in the two previous near identical strolls through the garden, namely choice a again.

Consider the following thought experiment; if we could change consciousness with other persons and given that the only change would be the subtle switch of phenomenal consciousness between me and Hitler and that in such a switch, one cannot bring anything other than the what it is likeness of being whatever person one is. Thus it is safe to assume that neither me, Hitler nor anyone around us would ever know what had happened. It also seems safe to assume that I would go on with Hitler's business as if nothing ever happened,

and he with mine. So unless I have some magical “soul” that can take credit for my moral “normalness” and Hitler has one that can be blamed for his “evilness”, then I think I would unfortunately- by moral luck, continue his practice of various crimes against humanity and he would be luckier and able to live a life without direct participation in such atrocities.

To illuminate what difference it makes to solve the problem of moral luck through hard incompatibilism, we can look at the thought experiment of the Christian or Islamic judgment day. The day where we will all have to answer for our sins and explain why we didn't behave better. On this day, if God is indeed omniscient, he must recognize that what is separating me from Hitler is the dumb luck of not having been born in a ruined economy to a mother who was mistreated by some particular Jews, not admitted into art school, marinated in national guilt/bitterness, having a thirst for power and seeing opportunity for it in peoples hate etc. etc. Keeping the libertarians out of it, if determinism about human actions is true then there are reasons – causes for Hitler's actions regardless of what they might be. So, if determinism about human actions is true, then God in the thought experiment would have to overlook those factors beyond our ultimate control before he judge us. And voila; stripped bare of all antecedent events leading up to the events of our lives- we are completely morally equal.

While the concept of God is highly irrelevant for the free will debate, the idea of having to answer for oneself at some point, in life or maybe after seems to have survived the religious secularization we have had in Europe/the west to a significant extent. Randolph Clarke discusses “heaven and hell- responsibility” in his paper called *On an argument for the impossibility of moral responsibility*. The argument he writes about is Galen Strawson's: “

- i) Nothing can be *causa sui*- nothing can be the cause of itself.
- ii) In order to be truly morally responsible for one's actions one would have to be *causa sui*, at least in some crucial mental respects.
- iii) Therefore nothing can be truly morally responsible. (Strawson 1994, 5)” (Clarke in French and Wettstein, p. 14)

This argument for the impossibility of moral responsibility is according to Clarke not very persuasive and the reason for this is that Strawson's understanding of freedom is so strict, that you cannot be free and then truly morally responsible if you have not somehow caused yourself. Why is this freedom and why do you need that strict a version of freedom for moral responsibility? Having chosen to argue in favor of hard incompatibilism in this thesis, I share Strawson's problems in answering these questions. They also happen to be the questions I tried to answer or at least explore in part one. In part that I also made an attempt to show

what freedom is from a compatibilist point of view and how moral responsibility can occur through such a concept of freedom (keeping the premise that we cannot have moral responsibility without freedom).

## 2.6. Criminal care

“Doesn’t there have to be *something* amiss in anybody who sets out knowingly to do evil? How shall we draw the line between exculpatory pathology of various sorts... and people who do evil “of their own free will”, knowing what they are doing? If we set the threshold too high, everybody gets off the hook; if we set it too low, we end up punishing scapegoats.” (Dennet 2003, p. 290).

First of all, knowingly doing evil “of their own free will” without a pathology of some kind is not something I think humans can do, for two reasons. My argumentation in part one brought me the conclusion that free will is not something humans can have, and therefore they cannot act out of it either. The second reason is because of our nature. In the above section I proposed that we are “moral” if there is not something wrong with us, and why that is so, from an evolutionary point of view. Following this logic, there is something “amiss in anybody who sets out knowingly to do evil”. The something might be a known pathology or something we do not yet understand, but it is reasonable to assume that a person would not be able to do evil if there was nothing wrong with him (in his mental life, physical health or that he had very harmful and deranging surroundings), on the basis of human nature. Drawing a line between exculpatory pathology and people who do evil of their own free will is therefore to draw an artificial line. But regardless of people having done harmful actions of their own free will or not, we as a society have to deal with them in some way. Enter the punishment versus rehabilitation debate.

Even though there may not be a difference between modern compatibilists and hard incompatibilists on how to deal with criminal individuals as our levels of expertise improve, I think there are reasons to favor the position that is not trying to defend the idea that we have moral responsibility for things we could not have done differently.

Derk Pereboom has a suggestion to an alternative way of handling criminals. In the sixth

chapter of his book *Living Without Free Will; Hard Incompatibilism and Criminal Behavior*, he divides the methods of dealing with criminals in two: rehabilitative and punitive. The punitive methods are the ones which by his understanding of hard incompatibilism needs to go, and this is also as he points out one of the reasons why more people do not endorse the position. His strategy is to dissect the punitive methods and look at what lies behind them, and then showing that they might not be so valuable after all, regardless of the truth or falsity of hard incompatibilism.

He mentions three standard reasons for justifying the punitive method and these are retribution, moral education and deterrence. According to Pereboom, hard incompatibilism rules out retribution because it is based on the idea that the agent who did something wrong deserves to be hurt in some way because he was morally responsible and has shown himself to be a morally despicable person. If hard incompatibilism is right, then no one can have the kind of ultimate responsibility for his or her actions that would justify the retributive punishment. The explanation Pereboom offer as to why this is a bad reason to punish criminals is that it does not have anything good intended like protecting society or morally improving the criminal. Also, that retribution in reality is not much better than vengeance, which more people would intuitively believe is wrong or at least very harmful for society.

Two objections can be made to this; one that retribution is different from vengeance because it invites closure while vengeance provokes more damage. Two, that retribution reinstates justice. Pereboom answers to these objections by suggesting that “in central cases, the sentiment of vengeance is an emotional expression of the sense of retributivist justice. In such cases, the sense of retributivist justice is explanatorily prior to the sentiment of vengeance, and thus if retribution has an independent justification, it might well not be threatened by the objection from vengeance. (-) But suppose it is unclear which prevails – retributivism or the objection. Then a case might be made that the plausibility of the objection makes it illegitimate to justify actual punishment policy retributivistically. Punishment – in particular, punishment designed to satisfy the retributivist goals – harms people.” And that “where there is a substantial likelihood that one’s justification for harming someone is illegitimate, then harming that person on the basis of that justification could well be morally wrong.” (Pereboom, 2001, p. 161).

When it comes to the purpose of morally educating through punishment Pereboom states that it is not clear that the punishment of criminals will serve to improve them (as might be

the case for children according to him). Also, as opposed to what might be the case with children, criminals more often than not *does* know that what they have done is unacceptable in society. “Indeed, without substantial empirical evidence that punishment can successfully educate criminals morally, it would be thoroughly wrong to punish criminals for the reason that it can realize this outcome. If one proposes to harm someone in order to achieve a salutary result, one must have very good evidence that harming him in this way can have the intended effect.” (Pereboom, 2001, p.163).

One argument often used in favor of punishment is that it can serve to give the criminal a chance to make amends; to right the wrong he has caused. This way he may have an easier time getting back into society- from which he has most likely felt excluded in many ways. Punishing someone is to show them respect for being rational, reliable and last but not least; a person with free will. However, as Pereboom points out, if the criminal offence is severe, for instance sexual abuse, violence or murder, then it is naïve to think that the wrong can ever be “righted”. To expect that justice have been restored after someone has served 2 years in (a Norwegian) prison for rape for example, is to underestimate the hurt that has been done to the victim.

Regarding the utilitarian purpose of punishment, which is to prevent future crimes by the criminal punished and other potential criminals lurking around by giving a clear message of what is acceptable and what is not in society, Pereboom suggests that this too can be problematic irrespective of hard incompatibilism being true or not. One of the reasons for this are the possible scenarios where the person who has committed the horrible crime is not found and the system might be tempted to frame and falsely convict the wrong person in order to gain the societal benefit of showing people that one cannot get away with serious crimes.

Further, Pereboom talks about self-defense as justification for punishment. As he points out, a lot of people can agree that it must at least be allowed to harm someone if it is for the protection of oneself or others in need of protection. There are however even problems with this motivation for punishment again regardless of the truth of hard incompatibilism. While it is ok to defend oneself or others from a criminal who poses an immediate threat, a criminal in custody (of the police, justice system etc.) does not normally pose such an immediate threat as he is being held under control. The self-defense would in this case be indirect. Indirect self-defense does not have the same justification direct self defense has.

The remaining reason Pereboom mentions as a justification for punishment is one that does not conflict with hard incompatibilism. This is detention as form of quarantine to protect society. Just as one would have to put up with being quarantined if one came back from a tropical vacation with a deadly and infectious disease, a dangerous criminal regardless of moral responsibility will have to be held apart from society in order to protect it. “One must note, however, that it is morally wrong to treat carriers of a disease more severely than is necessary to keep them from being dangerous to society. Thus, if the quarantine analogy provides our only justification for criminal detention, it will be morally wrong to treat those with violent criminal tendencies more harshly than is required to keep them from being dangerous to society. Furthermore, the less dangerous the disease, the less invasive the legitimate prevention methods would be, and similarly, the less dangerous the criminal, the less invasive the legitimate prevention methods would be.” (Pereboom, 2001, pp. 174-175). A problem with this justification for punishment is the idea that it may lead to the detention of innocent people who are viewed as potential criminals in the future. This is of course a massive human rights breach and would be morally problematic even though hard incompatibilism would not stand as firmly against it as it would against the other reasons for punishment Pereboom has presented. Pereboom's conclusion is clear; there are no persuasive reasons for punitive methods of dealing with criminals.

Rehabilitation is thus the best way to deal with criminals according to Pereboom and his hard incompatibilism, and as he argues, many justifications of dealing punitively with criminals are invalid independently of hard incompatibilism.

Pereboom's critique of these various approaches to dealing punitively with criminals are in my view interesting and useful when it comes to the development of a more and more humane way of handling wrongdoers, given that this development will continue. While, as Pereboom makes clear, it is not necessary to have a hard incompatibilist outlook on humans to feel like punitive reactions should become more seldom or disappear altogether, it does go very well with it.

## **2.7. Conclusion of part two**

In this part of my thesis I have made an attempt to establish what I have called a *naturalistic* understanding of morality. This involved making a distinction between *ultimate moral responsibility* and *practical moral responsibility*. I concluded that with the understanding of



freedom I have argued for in part one, we cannot have ultimate moral responsibility. However, with the naturalistic understanding of morality, there seem to be no obstacle for morality of the practical kind in the eyes of the hard incompatibilist.

This distinction means that it is not a matter of “all or nothing” with moral responsibility. While Strawson argues that even if hard determinism is true or believed to be so (for this matter hard determinism and hard incompatibilism are in the same boat, so they can be targeted with the same critique,) then that cannot change our reactive attitudes toward each other and ourselves because they are natural. I agree that our reactive attitudes would not change much, and because of the distinction I have made within morality they should not really be an issue.

A little bit of change would be in its place though. Not because of hard incompatibilism demanding it, but because our reactive attitudes, as Strawson describe them, adjust to our level of understanding of human behavior. I’m claiming that in the west in general and in Norway in particular, our reactive attitudes toward all levels of “troublemakers” and “wrongdoers” have adjusted with the growing expertise of psychology, biology, etc. With this as base, I contend that in the hypothetical case that it should be discovered and publicly believed that hard incompatibilism is true, this would only contribute to a revision and perhaps reform, but not a revolution of our reactive attitudes. This is what I think will happen regardless of hard incompatibilism, and I do not think hard incompatibilism would change much on its own.

I included a section where I discussed the problem of moral luck and how hard incompatibilism solves that problem. Also that hard incompatibilism in a way renders all humans morally equal in the “ultimate” way (not of course in the practical way), because of the lack of freedom.

The last section was one where I presented some of Perebooms views on why punitive reactions to criminals are unjustifiable from the point of view of incompatibilism, but also independently of this.

My view is that hard incompatibilism “compliments” the way society is hopefully moving anyways; toward more humanity as a result of higher levels of expertise on human behavior.

### 3. Thesis conclusion

My main question for this thesis was “can hard incompatibilism (with a naturalistic account of morality) offer a solution to the free will problem that doesn’t require a revolution of reactive attitudes? After this writing process, my answer to that is: yes it can.

I have given an account of why hard incompatibilism is my favored position in the free will debate. That consisted of an argument in favor of the agent causal libertarian definition of freedom in this context. A libertarian account of freedom is preferable to a compatibilist one because it fulfills the expectations of commonsense. The compatibilists versions of definitions of freedom seemed to me to violate the commonsense understanding and expectations of the term. The first of my “sub questions” was thus answered. What compatibilists, like Dennett for instance, call *freedom* is in my opinion better termed *control*. This amounts to voluntariness, autonomy, self-discipline and the like, and takes care of the worry that a belief in hard determinism or hard incompatibilism would pacify us and take away our possibilities of leading meaningful lives. I moved on to deliberate why I do not think humans can have agent causal libertarian freedom. My reason for not being convinced by this was that the idea that a free agent can cause his own actions in a special way, independent of prior events, seemed mysterious and not well accounted for.

The second “sub question” was “Why should moral responsibility have anything to do with freedom? - Why should we not simply hold that we do have freedom, but not the kind of ultimate moral responsibility that makes us condemnable?”. This question turned out to be trickier. The causal history principle shows how moral responsibility seems to be dependent on freedom. If we turn it the other way around though, I can only defend freedom’s attachment to ultimate moral responsibility through tradition and a weak intuition. There is enough agreement about the linking of freedom to moral responsibility across the board (among the contesting positions in the free will debate) however, that it can remain unaccounted for without causing me too much trouble.

As mentioned, I started out sympathizing with the position of hard determinism. As things progressed, I saw that hard incompatibilism could escape some of the problems hard determinism faced, and provide what I was after: a way of arguing against the practice of holding people morally responsible to the point of condemnation.

A distinction was made between ultimate moral responsibility and practical moral responsibility. I gave a naturalistic account of morality in order to show that morality need not be dependent on free will if we are talking about the practical kind. The ultimate kind of moral responsibility *is* arguably dependent on free will. The ability to be responsible was

described along with morality as an evolved character trait within our species as something useful for survival. This understanding of morality, added to the ability to act responsibly, is what I mean by the practical moral responsibility for which we can keep the reactive attitudes Strawson describe. The distinction let me keep the link between moral responsibility and free will, while denying freedom and the part about moral responsibility that didn't appeal to me, the ultimate kind, and retain the practical part.

Regarding the third “sub question”; “How might the negation of freedom and moral responsibility possibly benefit society and the practicalities of dealing with “bad” (criminal, violent, mean, cruel etc.) individuals?” I relied heavily on Pereboom.

When it comes to actions that go well beyond what is considered hurtful to others and can be described as “evil acts”, our reactive attitudes are less useful as a tool for interaction. In accord with Strawson, in these cases we modify our reactive attitudes according to how sane, sober, mature or reliable we consider the wrongdoer to be.

I agree with Pereboom that the justification for punitive ways of dealing with criminals suffers under the hard incompatibilist view. I also think, like Pereboom, that rehabilitation and/ or “quarantine” –detention of criminals in order to protect society, is better than punishment for the sake of hurting the wrongdoer. Comparing the situation today with the way criminals have been viewed and treated through history, I think it is safe to say that at least in Norway, the trend is moving toward a more and more humane way of dealing with criminals. This development can be seen as parallel with and a result of increasing knowledge about human behavior in areas like psychology, biology and neurology. The way I see it, hard incompatibilism and hard determinism are views that can compliment this development better than the others. This “complementing” is the possible benefit I expect that hard incompatibilism could have provided society with in the hypothetical scenario that it was to be publicly believed or proven correct.

Now, I am aware that I have not made a complete and waterproof case for my accounts of hard incompatibilism, practical moral responsibility and how society should deal with criminals. I hope however, to have given the reader some good reasons to consider hard incompatibilism as a worthy contestant of the free will debate, and a position that arguably has less theoretical problems than the rest of the positions, and not all the practical issues it is often accused of having. So where should we go from here? I think it would be incredibly interesting if there came about more experimental philosophy on the topic. More studies like

the two I have included here, about what people believe about philosophical questions before being exposed to philosophy courses. It will also be intriguing to pay attention to what goes on in the field on neurological research about consciousness, volition and generally about how our brains are wired. And most of all, I will be interested in seeing whether or not we continue to adjust our judgments according to newly acquired knowledge about human behavior.

## 4. Sources

### 4.1. Books:

- Churchland Paul M. *The Engine of Reason, the Seat of the Soul*, Massachusetts Institute of Technology, Massachusetts, 1995.
- Dennett Daniel C, *Freedom Evolves*, VIKING, New York, 2003
- Fischer John Martin, Kane Robert, Pereboom Derk, and Vargas Manuel, *Four Views on Free Will*, Blackwell Publishing, Oxford, 2007
- French Peter A., Wettstein Howard K., *Free Will and Moral Responsibility*, Blackwell Publishing, Oxford, 2005
- Kane Robert, *A Contemporary Introduction to Free Will*, New York- Oxford, Oxford University Press, 2005
- Mumford Stephen and Ajum Rani, *Getting Causes from Powers*, Oxford University Press, New York, 2011
- Nagel Thomas, “What is it like to be a bat?”, *Mortal Questions*, Cambridge University Press, Cambridge, 1979
- Nagel, Thomas, *Mortal Questions*, Cambridge University Press, Cambridge, 1979
- Pereboom Derk, *Living Without Free Will*, Cambridge University Press, Cambridge, 2001
- Pinker Steven, *The Better Angels of Our Nature: Why Violence Has Declined*, Penguin Group, London, 2012
- Tabensky Pedro Alexis, *Judging and understanding Essays on Free Will, Narrative, Meaning and the Ethical Limits of Condemnation*, Ashgate, Hampshire, 2006
- Van Inwagen Peter, *An Essay on Free Will*, Oxford: Oxford University Press,

Clarendon Press, 1983, p. 16.

- Wilkinson Richard, Kate Pickett, *The Spirit Level: Why More Equal Societies Almost Always Do Better* (/ *Why Greater Equality Makes Societies Stronger*), Bloomsbury Press, New York, 2009

## 4.2. Articles and electronic sources

- Bargh John A. and Morsella Ezequiel, *The Unconscious Mind*, Perspectives on Psychological Science, Association for Psychological Science, Yale University, 2008, [http://www.yale.edu/acmelab/articles/Bargh\\_Morsella\\_Unconscious\\_Mind.pdf](http://www.yale.edu/acmelab/articles/Bargh_Morsella_Unconscious_Mind.pdf)
- Bedke Matt, Deery Oisin, Nichols Shaun, Phenomenal Abilities: Incompatibilism and the Experience of Agency, In David Shoemaker (ed.), *Oxford Studies in Agency and Responsibility*. Oxford UP (forthcoming), 2012
- Bourget David, Chalmers David, What do Philosophers Believe?, Phil Papers, April 2013, URL= <http://philpapers.org/archive/BOUWDP.4.pdf>
- Brogaard Berit, D.M.Sci., Ph.D, Why The Mind May Be Physical It's not what you think, Published on November 27, 2012 in The Superhuman Mind
- Cholle Francis, What Is Intuition, And How Do We Use It?, Psychology Today, 2011, <http://www.psychologytoday.com/blog/the-intuitive-compass/201108/what-is-intuition-and-how-do-we-use-it>
- Crick Francis, Koch Christof, Consciousness and Neuroscience, 1997, (appeared in) *Cerebral Cortex*, 8:97-107, 1998, <http://www.klab.caltech.edu/~koch/crick-koch-cc-97.html>
- De Waal Frans: Moral behavior in animals, Filmed Nov 2011 • Posted Apr 2012 • TEDxPeachtree
- Denis, Lara, "Kant and Hume on Morality", *The Stanford Encyclopedia of Philosophy* (Fall 2012 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2012/entries/kant-hume-morality/>.
- Di Norcia, Vincent, *Ethics on the Brain*, Brains and Minds, Philosophy Now, Nov/Dec, 2012, [http://philosophynow.org/issues/87/Ethics\\_On\\_The\\_Brain](http://philosophynow.org/issues/87/Ethics_On_The_Brain)
- Duus-Otterstrøm, Gøran, "Betting Against Hard Determinism", *Res Publica*, 2008 (14): 219-235

- Fehr Ernst, Nowak Martin A, Sigmund Karl, “The Economics of Fair Play”, Scientific American Magazine, 2002, URL= [https://miside.uib.no/dotlrn/classes/det-historisk-filosofiske-fakultet/fil305/fil305-2013v/file-storage/download/105394531/Sigmund\\_%28The\\_economics\\_of\\_fair\\_.\\_.\\_.%29.pdf](https://miside.uib.no/dotlrn/classes/det-historisk-filosofiske-fakultet/fil305/fil305-2013v/file-storage/download/105394531/Sigmund_%28The_economics_of_fair_._._.%29.pdf)
- Garvey James, *Frank Jackson, latter day physicalist*, Posted by The Philosophers Magazine, November 12, 2012
- Greene Brian, Making Sense of String Theory, TED, 2005  
[http://www.ted.com/talks/brian\\_greene\\_on\\_string\\_theory.html](http://www.ted.com/talks/brian_greene_on_string_theory.html)
- Greenstreet, Stuart, *Free Will versus Natural Necessity?*, Philosophy Now, Nov/Dec 2012, [http://philosophynow.org/issues/93/Free\\_Will\\_versus\\_Natural\\_Necessity](http://philosophynow.org/issues/93/Free_Will_versus_Natural_Necessity)
- Güzeldere Güven, *Consciousness Resurrected*, Mind & Morals June/July, Philosophy Now, 2002, [http://philosophynow.org/issues/36/Consciousness\\_Resurrected](http://philosophynow.org/issues/36/Consciousness_Resurrected)
- Harris Sam, Discussion on Free Will, Dangerous ideas festival, Sidney Opera house, Sidney, 2013, URL= <http://www.youtube.com/watch?v=JM3raA1EwrI>
- Harris Sam, Life Without Free Will, Sam Harris The Blog, September 9. 2012, URL= <http://www.samharris.org/blog/item/life-without-free-will>
- Harris Tom, How Hypnosis Works, How Stuff Works, <http://science.howstuffworks.com/science-vs-myth/extrasensory-perceptions/hypnosis1.htm>
- Hofer, Carl, "Causal Determinism", *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2010/entries/determinism-causal/>.
- Jackson Frank, Epiphenomenal qualia, the philosophical quarterly, Vol. 32, No. 127, pp. 127-136, 1982, URL= <http://www.jstor.org/discover/10.2307/2960077?uid=3738744&uid=2&uid=4&sid=21102190418083>
- Langford, Michael, “Consciousness, Freewill and Language”, Mind and Brain Nov/Dec, 2011, *Philosophy Now*, URL= [http://philosophynow.org/issues/87/Consciousness\\_Freewill\\_and\\_Language](http://philosophynow.org/issues/87/Consciousness_Freewill_and_Language)
- Laplace, P., 1820, *Essai Philosophique sur les Probabilités* forming the introduction to his *Théorie Analytique des Probabilités*, Paris: V Courcier; repr. F.W. Truscott and F.L. Emory (trans.), A Philosophical Essay on Probabilities, New York: Dover, 1951

- Libet Benjamin, "Do We Have Free Will?", *The Journal of Consciousness Studies*, URL=  
<http://www.centenary.edu/attachments/philosophy/aizawa/courses/intros2009/libetjcs1999.pdf>
- Libet Experiments, The Information Philosopher, URL=  
[http://www.informationphilosopher.com/freedom/libet\\_experiments.html](http://www.informationphilosopher.com/freedom/libet_experiments.html)
- McKenna, Michael, "Compatibilism", *The Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, Edward N. Zalta (ed.), URL =  
<<http://plato.stanford.edu/archives/win2009/entries/compatibilism/>>.
- Nahmias Eddy, Morris Stephen, Nadelhoffer Thomas, and Turner Jason, "Surveying Freedom: Folk Intuitions about Free Will and Moral Responsibility", *Philosophical Psychology* Vol. 18, No. 5, October 2005, pp. 561–584,  
<http://www.brown.uk.com/brownlibrary/nahmias.pdf>
- Nelkin, Dana K., "Moral Luck", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL =  
<<http://plato.stanford.edu/archives/fall2008/entries/moral-luck/>>.
- Nichols Shaun, "The Folk Psychology of Free Will: Fits and Starts, Mind & Language", Vol. 19 No. 5 November 2004, Blackwell Publishing, Oxford, pp. 473–502.
- Nichols, S. 2007. "After Incompatibilism: A Naturalistic Defense of the Reactive Attitudes". *Philosophical Perspectives*, 21, 405-428.
- O'Connor, Timothy, "Free Will", *The Stanford Encyclopedia of Philosophy (Spring 2013 Edition)*, Edward N. Zalta (ed.), URL =  
<<http://plato.stanford.edu/archives/spr2013/entries/freewill/>>.
- Philosophy, Free will and politics, Jan 12th 2012, 18:19 by M.S,  
<http://www.economist.com/blogs/democracyinamerica/2012/01/philosophy>
- Randel Koons, Jeremy, *Is Hard Determinism a Form of Compatibilism?*, The Philosophical Forum, Volume XXXIII, No. 1, Spring 2002
- Rice, Hugh, "Fatalism", *The Stanford Encyclopedia of Philosophy (Winter 2010 Edition)*, Edward N. Zalta (ed.), URL =  
<<http://plato.stanford.edu/archives/win2010/entries/fatalism/>>.

- Robinson, Howard, "Dualism", *The Stanford Encyclopedia of Philosophy (Winter 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2011/entries/dualism/>.
- Savulescu Julian and Persson Ingmar, Moral Enhancement, *Philosophy Now*, Jul/Aug 2012, [http://philosophynow.org/issues/91/Moral\\_Enhancement](http://philosophynow.org/issues/91/Moral_Enhancement)
- Schwitzgebel Eric, "The Unreliability of Naïve Introspection", *Philosophical Review*, Vol. 117, No. 2, 2008, Cornell University, URL=<http://www.faculty.ucr.edu/~eschwitz/SchwitzPapers/Naive1.pdf>
- Smilansky Saul, Hard Determinism and Punishment: A Practical *Reductio*, *Law and Philosophy*, May 2011, Volume 30, *Issue 3*, pp 353-367, URL=<http://link.springer.com/article/10.1007%2Fs10982-011-9099-9?LI=true>
- Strawson Galen, The Impossibility of Moral Responsibility, *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, Vol. 75, No. 1/2, Free Will, Determinism, and Moral Responsibility Aug., 1994, pp.5-24, <https://philosophy.as.uky.edu/sites/default/files/The%20Impossibility%20of%20Moral%20Responsibility%20-%20Galen%20Strawson.pdf>
- Strawson, Peter F. (1962). Freedom and resentment. *Proceedings of the British Academy* 48:1-25.
- The History of the Free Will Problem Retrieved April 29, 2013, from Information Philosopher Web site <http://www.informationphilosopher.com/freedom/history/>
- Vihvelin, Kadri, "Arguments for Incompatibilism", *The Stanford Encyclopedia of Philosophy (Spring 2011 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/spr2011/entries/incompatibilism-arguments/>.
- Weed Laura, *Philosophy of Mind: an Overview*, *Brains and Minds*, Nov/ Dec, *Philosophy Now*, 2012, [http://philosophynow.org/issues/87/Philosophy\\_of\\_Mind\\_An\\_Overview](http://philosophynow.org/issues/87/Philosophy_of_Mind_An_Overview)
- Williams, Garrath, Praise and blame, *Internet Encyclopedia of Philosophy*, April 2006, <http://www.iep.utm.edu/praise/>