The mediation between brain imaging technologies and libertarian free will theories.

Master thesis Wisse Boomsluiter, s1365460 First supervisor: Ciano Aydin Second supervisor: Saskia Nagel

University of Twente Faculty of Behavioural, Management, and Social Sciences MSc Philosophy of Science, Technology and Society - PSTS Enschede, the Netherlands, January 2018 Contents

Abstract			3
1.	In	troduction	4
2.	Free will debate		7
	1.	Libertarianism	8
	2.	Indeterminism	10
	3.	Conclusions	11
3.	Agency and libertarian theories of free will		12
	1.	Agency	12
	2.	Agent-causal framework	13
	3.	Noncausal (volitionist) framework	15
	4.	Event-causal framework	16
	5.	Two-stage models of free will	17
	6.	Conclusions	18
4.	Br	ain imaging technologies and agency	18
	1.	fMRI	19
	2.	Libet experiments	21
	3.	Conclusions	24
5.	Technological mediation theory		24
	1.	Technological mediation	25
	2.	Directly perceiving agency	26
	3.	Limiting agency to the brain	26
	4.	Shift towards event-causal libertarian theories	29
	5.	Conclusions	30
6.	Conclusion		30
	1.	Discussion and follow-up research	33
	Bi	bliography	35

Abstract

The debate around free will is not new in philosophy, it has been discussed by a great deal of wellknown philosophers like Augustine, Descartes, and Kant. Free will is a topic that is debated about in contemporary literature, one of these contemporary discussions is within the field of brain imaging technologies (e.g. fMRI), as images derived from these technologies seem to have a bearing on a various amount of philosophical issues, including free will, consciousness, identity and agency. By making use of the technical mediation theory the author tries to answer how brain imaging technologies mediate accounts of free will. The free will debate holds many positions, the author of this thesis chose to focus on one particular position of free will: (metaphysical) libertarianism. In short, libertarians think the world is not determined and that agents have free will. As explained within this thesis libertarian theories of free will can be categorised within three sub-positions: (a) agent-causal; (b) noncausal (volitionist); and (c) event-causal.

The research question within this thesis is: "How do brain imaging technologies mediate libertarian accounts of free will?". The author of this thesis explains that the brain is given a pivotal role within libertarian theories, and that the brain is more often seen as the locus of the mind. Brain imaging technologies have strengthened this change of view in contemporary libertarian theories. Seeing the brain (and the self) as the locus of the mind also entails that the brain is seen as the locus of agency. Libertarian theories that endorse the idea that the brain is the locus of agency cannot fit into the agent-causal framework, and therefore make use of the event-causal framework. The author claims therefore that this mediation between brain imaging technologies and agency has strengthened a recent shift of libertarian theories from the agent-causal framework towards the event-causal framework. This shift itself might not even be noticed by the libertarian authors themselves, as the origin of the shift, the mediation of brain imaging technologies, is pushed to the background. This might be the most noticeable among libertarian accounts that have a relation to or make conclusions inferred from the measurements of brain imaging technologies.

1. Introduction

The debate around free will is not new in philosophy, it has been discussed by a great deal of wellknown philosophers like Augustine, Descartes, and Kant. Free will is also a topic that is debated about in contemporary literature (Timpe, n.d.), and there are many different views on what free will actually is, whether or not it exists, and whether it can co-exist with determinism and indeterminism. One of these contemporary discussions is within the field of brain imaging technologies, as images derived from these technologies seem to have a bearing on a various amount of philosophical issues, including free will, consciousness, identity and agency (Aydin in press-a). The free will debate is also central in ethical and moral considerations and is furthermore relevant to most other aspects of thought, including scientific reasoning. Furthermore, assuming that free will exists, how one thinks about the world and about certain concepts could possibly depend on his or her thoughts about whether the world is determined and whether one has a free will.

The free will debate holds many positions (hard determinism, compatibilism, hard incompatibilism, and libertarianism, with a large amount of sub-positions), to include all positions within this paper would be too complex and would make paper enormous size wise. I have therefore chosen to focus on only one position, namely libertarianism. In chapter 2 I will give a concise description of the various positions within the free will debate, as well as a more detailed description of libertarianism. For this introduction it is necessary to know that libertarians think we have free will, and that we do not live in a fully deterministic world. Furthermore libertarianism, one only has free will if one has alternative paths to chose from, and therefore could have done otherwise (Kane, 2007). Within libertarian theories, it is the agent that makes the (free) decision of which path to take.

The concept agency is of great importance to this thesis, because (as explained in greater detail in chapter 3) the concept of agency is central within libertarianism; this centrality is important to my research question. Therefore, before I continue, I will provide a basic description of agency. In most basic terms, an agent is a being that has the capacity to act; agency is therefore the manifestation of this capacity to act (Schlosser, 2015)¹. Because agency follows from action, most theories about agency are either derived or linked with theories about action. Within most theories of agency, actions are only attributed to agency when they are *intentional* actions. According to Mele (2003) a framework explaining agency should incorporate many kinds of (intentional) actions, including: a) rational as well as irrational actions, b) morally proper as well as morally improper actions, c) simple, routine actions as well agent involving actions, and (d) intentional mental actions (e.g. solving chess problems leading to deliberation) as well as overt intentional actions (involving peripheral bodily motion).

¹ The term agency as used in this thesis, should not be confused with "a sense of agency" often used in literature. The sense of agency requires the awareness of the performed action, while agency in this thesis involves only the manifestation of the capacity to act. Master thesis Page 4 of 38

Out of the various positions in the free will debate, I chose to focus on libertarianism. There are various reasons for why I chose to focus on libertarianism in this paper instead of one of the other free will positions. One of the reasons is, that libertarian theories often take an internalist perspective, which means that they see the origin for free will somewhere inside the person or agent. Within these theories (of which a few are discussed in the following paragraphs) the brain is often taken as the specific section where agency is performed and free will is realised. This internalist perspective, creates a certain connection between libertarianism and brain imaging technologies. As explained in detail in chapter 4 brain imaging technologies give us the perception of being able to look inside our brains. The connection between libertarianism and brain imaging technologies seems to be stronger for libertarianism than for the other positions in the free will debate.

One for the reasons for the connection with agency and the brain in libertarian theories is because proponents of contemporary libertarian theories of free will suggest that the brain might be partly indeterministic (e.g. Kane 2013; Doyle, 2013; Heisenberg, 2013). Libertarian Robert Kane (2013) for example claims that there is some indeterminacy within circulating impulses in feedback loops that happen in the brain. He explains that this indeterminacy within neuronal processes might complicate cognitive tasks, as it creates a kind of chaotic background. Kane further states that the question whether indeterminism is in certain places in nature (like the brain), is one that cannot be solved by philosophical reasoning alone, and that we need scientists (including neuroscientists) to make progress on this issue.

Doyle (2013) claims that the brain has evolved to the point where it can access quantum events, which he deems as an evolutionary advantage for the brain, as it makes freedom and creativity possible. Doyle further claims that "Free will is not an *ad hoc* development in humans [...]. It is a normal biologic property, not a gift of God or an inexplicable mystery." (p.240). As Doyle (2013) claims that new possibilities exist only within our minds, and neuroscientists might eventually be able to predict our thoughts and decisions if they could resolve the finest details of information storage in our brains. This is assuming that they have knowledge of all our past actions and all the randomly generated possibilities in the instant before our decision.

Neurobiologist Martin Heisenberg² (2009) claims that "There is plenty of evidence of chance at work in the brain: take the random opening and closing of ion channels in the neuronal membrane, or the miniature potentials of randomly discharging synaptic vesicles." (p.164-165). Furthermore, he states brains of animals continuously pre-activate, discard and reconfigure their options, as well as evaluate the short- and long-term consequences. However, he notes that there has been little investigation of the physiology of how this happens. While most of his research is about non-human animals (especially *Drosophila* species; often named fruit flies), in an article written in 2013 Heisenberg makes clear that non-human animals, as well as humans generate behaviour by themselves, and that therefore behaviour of most non-human animals as well as humans can be free (Heisenberg, 2013). What Heisenberg means with behaviour generated by

² Son of Werner Heisenberg, which was one of the key pioneers of quantum mechanics.

themselves (in his paper "her- or himself") is that the behaviour is not released by external stimuli, but is initiated from within the organism (Heisenberg, 2013).

Another reason for choosing libertarianism is that libertarianism is often overlooked in (contemporary) discussions about free will, as other positions are more popular within contemporary philosophy. I therefore think that more attention should be given to this position. Compatibilism is probably the most popular position within contemporary discussions about free will, despite that compatibilism is often quite hard to grasp for newcomers in the free will debate. That libertarianism is often overlooked in contemporary philosophy, is also problematic because I think the libertarian view is probably the most common (or popular) view lay people intuitively think of, and probably endorse, when asked about free will (a thought shared by Kane, 2007). That libertarianism is often overlooked becomes especially clear within contemporary discussions about free will. Within these discussions compatibilism is often seen as *the* way to reconcile free will with modern science, as there are also some libertarian proponents (e.g. Doyle, 2013; Kane, 2013) who have made attempts to do so (see also the examples given in the previous paragraphs), and the number has recently starting to increase3.

One prominent way to research brain functions in cognitive neuroscience is via brain imaging technologies (BITs). Brain imaging is a technique that provides measures of brain structure, as well as (depending on the technology) when and where certain aspects of brain activity occur (Gordon, 1999). According to Racine, Bar-Ilan, & Illes (2005), these neuroimaging technologies become increasingly pivotal in debates among scientists (e.g. Op de Beeck, 2010; Cox, & Savoy, 2003), and philosophers (e.g. Pauen, 2008), and are brought to the public's attention through media and lay publications. Furthermore, brain imaging technologies are increasingly used to correlate brain activation with psychological states and traits (Aydin, in press-a; Racine et al., 2005; for examples see: Pardo, & Patterson, 2013; Canli, et al., 2001). The connection between brain imaging technologies and the concept of agency will become important later on in the thesis (see chapter 4), as I will explain that I think that brain imaging technologies affect how we think about agency.

Because this thesis will be about the mediation between brain imaging technologies and libertarian theories of free will I will draw upon the technological mediation theory. According to this technological mediation theory, brain imaging technologies, like MRI scanners are not merely neutral intermediaries (Verbeek, 2015). Within the formation of many theories and interpretations that are about brain imaging technologies, or that rely on results of brain imaging technologies, the technology (e.g. fMRI) itself is often pushed to the background. Many of these authors do not take into account that brain imaging technologies themselves have an influence on their interpretations, conclusions and formation of theories. However, according to Verbeek (2015) how we perceive and understand certain concepts (e.g. the brain) cannot be understood without taking the mediating role of technologies (in this case brain imaging technologies) into account. The term mediation

³ Most probably because of the increasing interest in quantum mechanics, providing new insights on indeterminism within nature. Master thesis Page 6 of 38

within the technological mediation theory is used to describe "how [technologies] mediate the relation between humans and their world, amongst human beings, and between humans and technology itself" (Verbeek, 2005; p. 11). Technologies (like brain imaging technologies) are therefore not neutral tools to measure properties of humans and the world around us, but codetermine what they measure, and therefore frame concepts like the brain, brain processes, the self, and the world in a certain way (Aydin, in press-a). In short, the technological mediation theory is a framework that can be used to analyse the various types of relations between human beings, technologies and the world (Verbeek, 2015; Verbeek, 2016).

The research question of this thesis is: How do brain imaging technologies mediate libertarian accounts of free will?. This research question can be divided in various sub-questions, which are: (a) In what way is the concept agency central to libertarian theories of free will?; (b) How do brain imaging technologies affect the way we think about agency?; and (c) How does such an effect in the perception of the concept agency affect libertarian theories of free will?

In this thesis I investigate how brain imaging technologies mediate libertarian theories of free will. In order to understand this mediation, the argument is divided in three separate steps. First, I propose that the concept *agency* is central to libertarian theories of free will. Secondly, I propose that brain imaging technologies affect how we think about *agency*. Thirdly, I propose that such an effect in how the concept *agency* is perceived affects libertarian theories of free will. I should note that while I will touch many concepts throughout this thesis, the focus of this thesis is on brain imaging technologies, libertarianism, and agency and the mediation between these concepts. But before it is possible to look into the mediation between these concepts it is necessary to understand what these libertarian theories of free will are and where they are situated in the free will debate.

2. Free will debate

In order to understand how brain imaging technologies mediate libertarian theories of free will we first have to understand what libertarianism is. In this chapter I will provide a framework of what is meant with libertarianism, where libertarianism is located in the free will debate, and how libertarianism differs from other positions within the free will debate. Furthermore, in order to really understand libertarianism, I will explain the concept of indeterminism and why indeterminism is necessary for libertarianism.

The first step to gain an understanding of libertarianism is knowing where it situated in the free will debate. Most theories of free will can be categorised in four ways of thinking about free will and determinism: a) Hard determinism; b) Compatibilism; c) Hard incompatibilism; and d) (Metaphysical) Libertarianism. Within these four categories there are various sub-positions⁴. In this section I will provide a concise description of hard determinism, compatibilism, and hard incompatibilism. While the focus in this thesis will be on libertarianism, it is necessary to have a

⁴ While these four categories are agreed on by many philosophers, there are multiple different categorisations possible. Some authors of the sub-positions that I mention here (e.g. revisionism) claim that their position is in fact a different form/category of free will, which cannot be categorised within the position I propose. Master thesis

basic understanding of the other positions in the free will debate in order to understand how libertarianism differs from the other positions.

Hard determinism (often just called determinism or casual determinism as in Hoefer, 2003) is the idea that all events are necessarily caused by previous events and conditions set by laws of nature (Hoefer, 2003). Because everything in the world is pre-determined free will is not possible. Hard determinism entails that there is only one timeline, there are no possible alternative futures. All previous events were building blocks for the present that is now, and these past events are the cause of the events yet to come (James, 2009). What we call the present is only compatible with clear path for the future events, and there are no possible alternative futures (James, 2009).

Compatibilism (or soft determinism as in James, 2009) is the idea that free will is compatible with determinism. According to compatibilism a deterministic world does not necessarily exclude free will. Compatibilists think that even if all the choices and actions agents make are determined, the agents can still make choices freely (Timpe, n.d.). There are many compatibilistic accounts that give an explanation to how agents can have free will in a deterministic world. These various accounts involve refuting certain premises of the Classical Incompatibilist Argument (as described in McKenna, 2004).

Hard incompatibilism (or pessimism as in Timpe, n.d.) is similar to hard determinism as both viewpoints maintain that free will does mot exist and that determinism is incompatible with free will. However hard incompatibilists go a step further by claiming that not only determinism, but also indeterminism is incompatible with free will. Contemporary advocates of hard incompatibilism are agnostic about the truth of free will or determinism (Fischer, Kane, Pereboom, Vargas, 2007; see also Pereboom, 2007). They state that we do not have enough knowledge about free will, determinism and indeterminism. There are various sub-forms of hard incompatibilism like illusionists, who claim that free will itself is an illusion (Smilansky, 2001), and revisionists, who claim that we should revise our common sense notion of free will and moral responsibility (Fischer et al., 2007).

Libertarianism

Now to turn to (metaphysical⁵) libertarianism. Libertarians think that free will cannot exist in a completely determined world. A libertarian free will seems therefore to require indeterminism of some sort, opening up the possibility of alternative paths to choose from (Kane, 2007). I would like to note that throughout my paper I frequently refer to works of Robert Kane. The reason for this is that Robert Kane is a prominent figure in (contemporary) libertarian theories of free will⁶. Libertarians do think that free will exists, and that it is therefore possible for an agent to have made a different choice at a particular time. Libertarian explanations of how free will makes it possible for agents to make different choices vary; I will describe these explanations in more detail in

⁶ According to Doyle (2011e) Robert Kane is acknowledged as the leading spokesman for Libertarianism. Master thesis

⁵ To avoid confusion with the political use of the term libertarianism some authors have decided to use the pre-fix metaphysical.

chapter 3. An often heard critique of libertarianism is, that, if free will is not compatible with determinism, it seems neither to be compatible with indeterminism. It seems not compatible with indeterminism because an undetermined event is related to chance; it occurs spontaneously and is not controlled by anything or anyone (Kane, 2007; Pereboom, 2007). There are different libertarian positions (described in detail in chapter 3) that try to explain how a libertarian free will is possible, including agent-causal, noncausal (volitionist) and event-causal explanations (see fig. 1 for a schematic overview of the free will positions as used in this paper).



Fig. 1. Schematic overview of free will positions as used in this paper.

An often used term to refer to libertarianism is incompatibilism⁷ (e.g. Clarke & Capes, 2017). Incompatibilism refers to the idea that determinism is incompatible with free will, which is an idea that also holds true within libertarianism. There are however a few distinctions to be made between incompatibilism and libertarianism. First of all, incompatibilism includes libertarian as well as hard deterministic theories (Pérez Otero, 2016), as both sets of theories have in common that free will is incompatible with determinism, thus that one cannot have free will in a deterministic world. Hard deterministic theories, as opposed to libertarian theories, refute indeterminism and maintain that humans have no free will, because we live in a deterministic world (Hoefer, 2003). Libertarian theories are therefore only a subset of all incompatibilist theories.

Furthermore, while incompatibilism and compatibilism are certainly opposites, libertarianism and compatibilism are not. Compatibilistic theories, imply that free will and determinism are compatible, and therefore that free will is possible in a deterministic world. However, stating that free will is possible in a deterministic world, does not necessarily mean that free will is impossible in an indeterministic world. From a compatibilistic viewpoint one could also

 ⁷ This is most likely done to avoid confusion with the now more common political use of the term libertarianism, a more correct alternative that could be used is "free-willism".
 Master thesis
 Page 9 of 38

state that it does not matter whether or not we live in a deterministic or indeterministic world, because in either worlds we have free will.

As said before, libertarianism requires indeterminism of some sort. This indeterminism is necessary in order for an agent to have the possibility to act otherwise given the same circumstances. In order to fully understand how libertarians think it is possible for agents to have free will, it is necessary to understand the relation between indeterminism and libertarianism.

Indeterminism

Getting to know what libertarians mean with indeterminism is necessary in order to get a better understanding of libertarianism. Libertarian proponents all include some kind of indeterminism in their theories because they think that free will is not possible in a deterministic world. As explained before, determinism is the idea that all events are necessarily caused by previous events and conditions set by laws of nature; therefore, there is only one possible future, and this future has already been set-in-stone at the beginning of the causal chain of events. Indeterminism is often seen as the opposite of determinism, however, this is only partly true. It is true in the sense that an indeterministic world the future is not yet decided for, and that events are not necessarily bound by the laws of nature. Indeterminism also entails indeterministically caused events, which may mean uncaused (or self-caused) events, as well as probabilistically or randomly caused events, opening up alternative possibilities for certain events. It is however not true that indeterminism necessarily means that everything (or every event/action) is indeterministically caused (or uncaused), nor that there are no natural laws at all; indeterminism only entails that some events or actions are indeterministically caused or uncaused (Popper, 1950). Thus, in an (partly) indeterministic world events are not (necessarily) caused by previous events, nor are they (necessarily) bound by laws of nature, instead they are related (not synonymous) to randomness, or chance.

According to libertarian accounts of free will, indeterminism is necessary for free will, but also for agency, as it allows persons to choose from alternative paths in the future (Pérez Otero, 2016; Kane, 2007). By being able to make this choice, one acquires agency, as the agent self, and not something else determines his or her actions (Pérez Otero, 2016; Kane, 2007). It is important to note that libertarian proponents do not all think that every action is done from our own free will, therefore not all actions require agency of the person. It is often enough that only some actions are done of our own free will (Balaguer, 2004; Kane, 2007). According to Kane (2007) only will-setting or self-forming actions are necessary for free will. These will-setting or self-forming actions are the actions that makes a person the way he or she is, and are required for ultimate responsibility, which Kane sees as a core aspect of a libertarian free will. Although I will come back later (chapter 5) on self-formation, I should note that this is not the focus of this thesis. Neither is it my intention to discuss the concept of responsibility in this thesis. The focus of this thesis is on the concepts agency, brain imaging technologies and libertarianism. The combination of indeterminism and probability makes it possible for agents to make a different choice given the exact same past (including prior mental and physical events) right up to the moment before the choice happens (Balaguer, 2004; Kane 2007;), or in some accounts of libertarianism before the decision-making

process (or deliberation) starts. Furthermore, while these indeterministic alternatives paths seem to be necessary for free will, they are not sufficient on their own, as one can think of examples in which one may have alternative possibilities but the agent has no control over the actions that create these possibilities (e.g. missing a golf put because of a twitch), therefore these actions are not performed of their free will⁸ (Kane, 2007).

These alternative paths thus open up the possibility of different possible futures given the exact same past, including the psychological and physical history of the agent up to that moment. One might now get the feeling that these so seemingly undetermined free choices are in reality not so free, as for a person to make a different choice given the same deliberation would seem irrational or arbitrary (Balaguer, 2004; Kane, 2007). This is a valid concern for libertarian theories of free will, and the different libertarian positions have different solutions to this problem. One way this problem gets tackled is to introduce new forms of causation, like agent-causation, by which free actions are not determined events occurring prior to the action, but these free actions do also not occur merely randomly or by chance (Balaguer, 2004), instead they are directly caused by the agent (Kane, 2007). These new forms of causation are seen as a form of substance-causation (see agent-causal framework in chapter 3). In different words, according to libertarian theories the indeterministic choice one can make is not necessarily random, but indeterministically caused (or non-random uncaused as in Balaguer, 2004), because the choice is willed (and therefore caused) by the agent (Balaguer, 2004; Kane 2007). That the choice is willed, entails that it is done for certain reasons that the agent endorses, which is why it is done on purpose instead of accidentally (Kane, 2007). Indeterminism does not necessarily rule out causation nor does indeterministic causation mean that it is happening only because of luck or chance (Pérez Otero, 2016; Kane, 2007). Indeterminism is consistent with nondeterministic and probabilistic causation, it only rules out deterministic causation: causation where the outcome is unavoidable (Kane, 2007).

Conclusions

Within this chapter I gave an overview on what libertarianism is and where libertarianism is situated in the free will debate. Furthermore, I explained how libertarianism differs from other positions, and how the relevant terminology should (and should not) be interpreted and used. At last I tried to give a basic understanding of one of the core concepts of libertarianism, namely indeterminism, as it is this concept that really differentiates libertarianism from the other free will positions. It should now be clear that according to libertarianism, one only has free will if one has alternative paths to chose from, that these alternative paths require some form of indeterminism, and that it is the agent that makes the (free) decision of which path to take. A libertarian free will therefore seems to require agency. But what exactly is meant with agency? Furthermore, how are agency and libertarianism exactly related? In the next section I will give a detailed description of agency, the sub-positions of libertarianism (see also fig. 1), and explain the relations between agency and libertarianism. I will furthermore propose that the concept of agency is central to

 ⁸ Even though these actions are done by accident, they are still done by these agents, the agents are therefore still responsible (Kane, 2007).
 Master thesis
 Page 11 of 38

libertarian theories, because both agency and the different positions within libertarianism use the same metaphysical frameworks.

3. Agency and libertarian theories of free will

In order to answer the question how brain imaging technologies mediate libertarian accounts of free will, it is first necessary to establish if and how the concept of agency is central to libertarian theories. The reason being, that I propose later on (chapter 5) that the concept of agency is the bridge in understanding how brain imaging technologies mediate libertarian accounts of free will. It is therefore necessary to gain a good understanding of what agency entails, and how agency is related to libertarian theories of free will. Furthermore, I will give an overview of the libertarian sub-positions. These libertarian sub-positions are relevant to the research question as I propose in chapter 5 that there is a recent shift visible from one libertarian sub-position towards the other.

I will start this chapter by describing agency. I will do this by explaining what the standard conception and the standard theory of *action* are, and how they translate to a standard conception and standard theory of *agency*. After that I will explain the three metaphysical frameworks of agency (event-causal, agent-causal and noncausal), and show how libertarian theories are connected to these frameworks. At last I will go into detail in one specific contemporary eventcausal theory, namely two-stage models of free will.

Agency

An agent is a being that has the capacity to act; agency is therefore the manifestation of this capacity to act (Schlosser, 2015). In this definition a being can refer to a human being, a nonhuman animal and in some fields (such as Science and Technology Studies) artefacts can also perform agency. This chapter will be limited to only accounts of human agency as free will theories are most often focused on humans only. Furthermore, I should note that the term agency as used in this thesis, should not be confused with "a sense of agency" often used in literature (e.g. Gallagher, 2007). The sense of agency requires the awareness of the performed action, while agency in this thesis requires merely the manifestation of the capacity to act. Within the philosophy of action, a standard *conception* of action as well as a standard *theory* of action are formed, and from this it is possible to obtain a standard conception of agency and a standard theory of agency (Schlosser, 2015)9.

The standard *conception* and the standard *theory* of action are similar in the way that they view agency as an intentional action, however there are also some differences. According to the standard *conception* of action exercising agency can be seen as performing an intentional action, and in many cases unintentional actions that derive from intentional actions (Schlosser, 2015). However, according to the standard *theory* of action the exercise of agency can be seen as the instantiation of the correct causal relations between (mental) states and events in which the agent

⁹ It is possible to obtain a standard conception of agency and a standard theory of agency, from the standard conception of action and the standard theory of action because, as explained in the beginning of this section, agency is the manifestation of the capacity to act. Master thesis Page 12 of 38 is involved (Schlosser, 2015). According to the standard *conception* of action one needs to be have the capacity to perform intentional acts, while according to the standard *theory* of action one only needs to have the correct functional organisation (Schlosser, 2015).

The idea that agency should be about intentional acts is support by Alfred Mele. According to Mele (2003) (a major part of) a theory of (human) agency should provide a framework to explain intentional actions. According to Mele this framework should incorporate many kinds of actions, including: a) rational as well as irrational actions, b) morally proper as well as morally improper actions, c) simple, routine actions as well agent involving actions, and (d) intentional mental actions (e.g. solving chess problems leading to deliberation) as well as overt intentional actions (involving peripheral bodily motion).

The standard *theory* of agency belongs to the event-causal framework of agency, which is one of the three main metaphysical frameworks of agency, the others being the agent-causal approach and the noncausal (volitionist) approach (Schlosser, 2015). The last two approaches also include kinds of agency that do not fit in the standard *conception* of agency or the standard *theory* of agency. The pivotally of the concept of agency to libertarian theories should now become apparent, because the categorisation of the different libertarian theories is often done via the three metaphysical frameworks of agency (e.g. Clarke & Capes, 2017; Balaguer, 2004). Later on in the thesis (see chapter 5) I will claim that within libertarian theories there is a recent shift from the event-causal framework towards the agent-causal framework. In order to understand how this shift came about, it is necessary to have a good understanding of these different frameworks. In the next sections I will therefore explain the metaphysical frameworks and how libertarian theories fit within these frameworks, starting with the agent-causal framework.

Agent-causal framework

As explained before, the agent-causal framework is one of the three main metaphysical frameworks of agency. While the event-causal framework is the most widely accepted framework in contemporary philosophy, the agent-causal framework is more popular among contemporary libertarian theories¹⁰ (Kane, 2007). The agent-causal framework provides a non-reductive account of agency, as the agent's role is construed in the exercise of irreducible agent causation (Schlosser, 2015; Pereboom, 2007; O'Connor, 2000). The event-causal framework on the other hand can be seen as a reductionist framework, as the agent gets reduced to mental states or event (Schlosser, 2015). Proponents of agent-causal theories disagree with the standard theory of agency, as they argue that initiating action should not be reduced to the capacity to act intentionally, or because of reasons. Instead they argue that the exercise of agency could happen spontaneously, for no reason and that the agent does not need to have any prior intent; while reasons and intentions might influence how an agent acts, the source of the agency lies in the power to initiate (Schlosser, 2015; Pereboom, 2007). In this framework intentional actions are caused, but mental states and events

¹⁰ The event-causal framework is popular among compatibilistic theories of free will. The reason that the agent-causal framework is less used in contemporary philosophy than the event-causal framework, even though it is more popular within libertarianism, is because there are more compatibilists than libertarians.

are neither in the causation nor in the explanation of these actions (Mele, 2003). Within the agentcausal framework the agent is seen as some kind of persisting substance, and causation by an agent can thus be seen as causation by a substance (Clarke & Capes, 2017; Schlosser, 2015; Kane, 2007). Because a substance cannot itself be an effect, the agent is the *causa sui* (cause of itself) of its own (free) decisions (Clarke & Capes, 2017; Kane, 2007). In these theories substance is used to refer to self-standing entities; entities that instantiate properties but are not instantiated themselves (Whittle, 2016). In other words, agency is seen as the most basic element, and cannot be reduced to other elements (like mental states or events). Agent-causality thus proposes that events are not the only possible "entities" that can cause (other) events, but that agents can also cause events to happen (Markosian, 1999).

Because this framework might be considered complex or even superficial, it might be helpful to compare this framework with behaviourism¹¹, an approach that has many similarities with this agent-causal framework. Behaviourism is a psychological approach to understand behaviour of humans and animals. Proponents state that we should not rely on human introspection to explain why people act in certain ways, but rather explain behaviour: (a) as a response to certain stimuli that can stimulate reflexes (classic conditioning), or (b) as the of a history of reinforcing or punishing events (operant conditioning). Proponents of this approach therefore black-box the mental and internal processes of behaviour, either stating that these processes do not exist, or that they might exist but are not necessary to explain how behaviour occurs. This is interesting in comparison with the agent-causal framework, as one might think that the agent-causal framework can be seen as black-boxing the mental aspects of agency as well. The agent-causal framework entails that it is the agent that causes the actions of the agent, not (limited to) the mental events that might be occurring at that time, in the same way as behaviourism would state that it is the human or animal that causes their actions, not (limited to) the mental processes that might be occurring at that time. Note however that behaviourism and the agent-causal framework are not the same, and this comparison should only be used in order to get a better understanding of the agent-casual framework.

Within agent-causal theories a free decision is caused by the agent. Within agent-causal *libertarian* theories this causation by the agent (what the agent causes, as well as the process of causation) should not be (causally) determined by events happening prior to this agent-causation (Clarke & Capes, 2017; Kane, 2007; Pereboom, 2007). One might think that it is improbable, or even impossible, for agent-causal framework to work within non-libertarian free will positions, and therefore it might seem redundant to talk about agent-causal *libertarian* theories, as opposed to just agent-causal theories of free will. However, this is not necessary the case, as Markosian (1999) shows us, it is possible to use the agent-causal framework with event-determinism¹². Markosian explains a so called "Frankfurt example" to explain how the agent-causal framework can be used for compatibilism accounts. The example Markosian uses states that there are two persons (Franny

¹¹ The best known behaviourists are most likely Burrhus Frederic Skinner (operant conditioning) and Ivan Petrovich Pavlov (classic conditioning).

¹² Every event that occurs is made physically necessary by previous events (Markosian, 1999; p. 9). Master thesis
Page 14 of 38

and Zoe) who are planning to assassinate a third person. Franny is worried that Zoe will back out the assassination at the last moment. To prevent this from happening Franny places a braincontrol device in Zoe's head, that will force Zoe to pull the trigger *only* when Zoe decides to back out. Zoe however does not back out, and the device never "forced" Zoe to change her action. From a compatibilistic perspective the decision is made freely as Zoe was in this case not forced to do otherwise, and her actions confirm her intentions. The decision would however not be free from a compatibilistic perspective, if she had decided not to go through with the assassination, as the device would have forced her to change her action. From a libertarian perspective, no matter if the device kicked in, the choice was not made freely because Zoe could not have done otherwise, as the action was determined by events happening prior to what seemed to be a "free decision". In other words, from a libertarian perspective Zoe was "tricked" into thinking her choice was made freely.

Noncausal (volitionist) framework

The noncausal framework is not as popular as the other two frameworks in contemporary philosophy. However, I think it is worthwhile to explain this framework, as knowing what is excluded from the other two frameworks helps in gaining a better understanding of the other two frameworks. Within the noncausal framework, also called the volitionist framework, agency is performed by volitions, which are seen as acts of the will. Intentional actions are intrinsically formed by a basic mental action, a volition (Clarke & Capes, 2017; Schlosser, 2015). These volitions in noncausal theories (as opposed to mental states in event-causal theories) are themselves entirely uncaused, and only related to intrinsic properties (they are *sui generis* acts) (Clarke & Capes, 2017; Schlosser, 2015). Therefore, if an agent decides to raise an arm (an intentional action), this agent first has the intention to decide to raise that arm (the basic mental action, or volition). This event is therefore not (pre-)determined, but caused by the person's intention. This also means that there is no agency assigned to the (external) intentional actions, but to the (internal) intention itself (the volition).

The noncausal framework only takes intentional acts into account, as other acts are not viewed as an exercise of the person's agency (Clarke & Capes, 2017). If the act is not intentional, the person executing the act could not have willed it (whether the result of the action was positive or negative does not matter), which means that the basic mental action (the volition) was absent.

There are (at least) two different perspectives noncausal libertarian theories take to situate the noncausality within the framework. The first one is that intentional actions, or at least some of them (like free actions) are uncaused, and have a noncausal explanation involving mental events or states (Mele, 2003). The second one differs in that all intentional actions may be caused, but that explanations in terms of beliefs and desires (forming in the unconscious) are uncaused explanations (Mele, 2003). In other words, noncausal libertarian theories either require that free actions are uncaused (they are a *causa sui*), or that they are indeterministically caused.

Event-causal framework

In comparison with the noncausal and agent-causal framework, the event-causal framework is the most widely accepted view in contemporary philosophy (Schlosser, 2015). Next to libertarian accounts, the event-causal framework is often used in compatibilistic accounts of free action. Mele (2003) provides us with a pair of theses for an event-causal perspective of intentional action. The first being that intentional actions are (deterministically or indeterministically) caused. The second one being that for any intentional action it is in principle possible to give a causal explanation partly framed in terms of mental events or states. Event-causal theories are based on the idea that there is a relation between the mental states (beliefs, emotions, desires, etc.) of the agent and events involving this agent. Agency is thereby assigned to the event-causal relations between these agent-involving states and events (Schlosser, 2015; Pereboom, 2007). According to the eventcausal framework, actions can always be reduced to certain events, however, events are only actions when there is a causal relation between the event and the mental states of the agent (Schlosser, 2015). In this case free actions are caused by the mental states of the agent. Within the event-causal framework mental states are caused by other events (mental states are also seen as events). This is opposed to both agent-causal theories and noncausal theories, where actions or mental states are not caused by events but are either uncaused, indeterministically caused (see noncausal framework) or caused by a substance (see agent-causal framework).

Libertarian event-causal theories differ from compatibilist accounts in the sense that libertarian accounts require that *some* of the events that cause actions are indeterministically caused (Clarke & Capes, 2017; Pereboom, 2007). Actions are performed because of certain reasons, and, until the action is actually performed, there remains a chance that the agent will not perform the action (or, according to some theories, a different action is performed instead) (Clarke & Capes, 2017). In other words, event-causal libertarian theories require that there is a possibility for the agent to do otherwise. Agency within event-causal libertarian theories is therefore more complex in comparison to other accounts using the same framework. As one can say agency is not merely assigned the event-causal relations between these agent-involving states and events but persons are also given agency with the possibility to stop the action, or to substitute it for a different action. However as pointed out by Clarke & Capes (2017) it is also possible to state the opposite, that by allowing indeterminism to create alternative possibilities agency is reduced instead of enhanced. One could for example state that because the alternative possibilities are indeterministic created, the person has no influence over which of the possible outcomes will be selected, thus undermining the responsibility of the agent, instead of enhancing it.

There are also free will theories that use the event-causal framework that can be situated inbetween libertarian and compatibilist theories. These theories, called two-stage models of free will, take (a limited form of) determinism as well as indeterminism as necessary factors for free will. I will describe these two-stage models of free will in the next section. The reason that I include this theory is that it will be used as a main example later on (chapter 5) in this thesis to show that the proposed shift from the agent-causal framework towards the event-causal framework might go unnoticed. Furthermore, I think that this theory is a worthwhile to mention, as it combines certain aspects of the free will debate in innovate ways, and might invoke new insights in the free will debates.

Two-stage models of free will

Before I discuss two-stage models I should note that the information about the two-stage models that I describe in my thesis originates mostly from Doyle. While there are other authors that have described their own two-stage model, or theories that fit into the category of two-stage models as described by Doyle, they do not go into detail about the concept of a two-stage model in general.

As mentioned before two-stage models of free will are models that combine a limited form of determinism as well as indeterminism (sometimes called comprehensive compatibilism, as in Doyle, 2011b). Because these models require indeterminism, they do seem to fit within the libertarian position of free will (which I will come back to later on in this section). While this view sees determinism as necessary, it rejects the view of pre-determinism (Doyle, 2011b), which is the idea that there is one causal chain of events that go back to the origin. Determinism is in this sense limited to adequate determinism, or statistical determinism, which incorporates quantum mechanics, but at the same time claims that the (indeterministic) uncertainty that comes from quantum mechanics is negligible (Doyle, 2011b). Indeterminism on the other hand is limited to the generation of alternative possibilities for action.

These models propose that making a decision is a process, instead of limiting the decision making to a single moment. Prior to this decision process there is the fixed past, in the sense that one cannot change past events. The first stage in this process is the generation of alternative possibilities, which are at least partly generated indeterministically (e.g. by quantum mechanics), this opens up the possibility of the agent having acted otherwise with the same circumstances (Doyle, 2011c). The second stage in the process is to evaluate and select these alternatives, which happens through deterministic deliberation, meaning that the selection is determined by our character, our values, our motives, and our feelings (Doyle, 2011c). In order to claim that our actions are caused by ourselves, they need to be at least somewhat determined, however this should not mean that our actions are pre-determined from (long) before even starting the process of making a decision, therefore indeterminism is required to open up alternative possibilities. However, in order for this decision process to be considered intelligible these indeterministic alternative possibilities cannot be the direct cause of our actions, therefore it is required that the alternative possibilities are generated first, and that the direct cause of our actions (the evaluation of alternatives) is causally determined.

Compatibilist and libertarians agree on the idea that humans¹³ have free will. However, they disagree as to whether such a free will can be compatible with determinism and whether it (necessarily) requires indeterminism of some sort. The free will proposed in these two-stage models is compatible with determinism as well as indeterminism. It could therefore be seen as a combination of a compatibilistic and a libertarian free will. Authors of two-stage models assume that indeterminism is necessary for our will to be free, but at the same time claim that determinism

is necessary for this free will to be intelligible. It follows the libertarian conception of free will, that in order to be free one could have done otherwise under the same circumstances. Doyle (2013) defines free will as: "... a two-stage creative process in which a human or higher animal freely generates alternative possibilities, some caused by prior events, some uncaused, following which the possibilities are evaluated and one is "willed," i.e., selected or chosen for adequately determined reasons, motives, or desires." (p.243).

Some authors (e.g. Doyle, 2013; Heisenberg, 2009) of these two-stage models of free will claim that the processes described in these theories can be situated inside our brains. As Doyle (2013) points out, new possibilities (referring to the alternative paths as discussed before) exists only within our minds, and neuroscientists might eventually be able to predict our thoughts and decisions if they could resolve the finest details of information storage in our brains (given knowledge of all our past actions and all the randomly generated possibilities in the instant before our decision).

Conclusions

Within this chapter I described the concept of agency and established that agency is central to libertarian theories by pointing out that libertarian theories are categorised by the three metaphysical frameworks, namely the event-causal, agent-causal, and noncausal frameworks. I furthermore gave an overview of these metaphysical frameworks and explained how libertarian theories fit within these frameworks. Establishing that agency is central to libertarian theories is only the first step necessary to answer the research question. I propose later on (chapter 5) that brain imaging technologies mediate libertarian theories of free will, because brain imaging technologies mediate the concept of agency. This however introduces a few questions. What exactly are brain imaging technologies, and how do they work? What is the relation between brain imaging technologies and agency? How do brain imaging technologies mediate the concept of one particular brain imaging technology (fMRI), and by thoroughly exploring the Libet experiments, which are famous experiments concerning both free will, agency and brain imaging technologies.

4. Brain imaging technologies and agency

In the previous chapter I tried to make clear how the concept of agency is central to libertarian theories of free will. The next step to show how brain imaging technologies mediate libertarian accounts of free will, is to show how brain imaging technologies affect the way we think about and perceive agency. This will show what the role of brain imaging technologies is within this mediation process. Later on in this thesis (chapter 5), I will show how this impact on agency by brain imaging technology affects libertarian theories of free will.

This chapter will be mainly about brain imaging technologies, which, according to Gordon (1999) provide *unambiguous* measures of brain structure, as well as (depending on the technology) when and where certain aspects of brain activity occur. I will however show that the measures of brain imaging technologies are not as unambiguous as they initially seem. I will show this by going

Master thesis

into detail about one particular brain imaging technology (fMRI), and use this technology to show it might be wise to question the validity of conclusions and interpretations made by authors (including philosophers) that follow from brain imaging technologies. After that I will go into detail about probably the most known experiments surrounding brain imaging technologies and free will, the Libet experiments. I will use the Libet experiments to show how brain imaging technologies can impact the concept agency. The section will furthermore give a practical example of how brain imaging technologies are used within neuroscience and philosophy, and will be used to support claims made in the first section of this chapter.

fMRI

In this chapter I will focus on one particular brain imaging technology, namely functional magnetic resonance imaging (in short fMRI). Because I want to show that brain imaging technologies like fMRI mediate our perception of agency, it is necessary to understand what is meant with a brain imaging technology like fMRI, how such a brain imaging technology works, and how we should interpret conclusions made from brain imaging technologies like fMRI. The reason that fMRI is chosen as opposed to other brain imaging technologies is that fMRI is seen as one of the most advanced brain imaging technology and is the most used brain imaging technology within contemporary (neuro-)science as well as within contemporary philosophical reasoning.

fMRI is seen as a non-invasive method to measure neuronal activity in the human brain (Heeger & Ress, 2002). The development of fMRI prompted a new subfield within neuroscience (Roskies, 2007), and many studies about fMRI have been picked up by the media and lay publications (Racine et al., 2005). The results of fMRI research are presented as greyscale images of brains where the degree of neural activity (brain activity) is marked with various colours (Roskies, 2007; Klein, 2010). This easy accessible visual way of presenting non-visual neural information is problematic for lay people, as well as scientists, as they might lack the necessary information and form conclusions without fully understanding what the images are describing in certain settings (Roskies, 2007). Racine et al., (2005) performed a meta-analysis of news articles concerning fMRI research. Within this meta-analysis they found that most of the news articles (67%) provide no explanation at all about the capabilities and limitations of fMRI. While it can be said that 33% did provide some explanation about the capabilities and limitations of fMRI, it is not clear whether this is enough to fully educate the reader of these articles. Most news articles therefore fail to provide lay people the necessary information to fully understand the results from fMRI research. Furthermore, articles and news report about fMRI results are often presented in a loaded way (e.g. "fMRI knows your secrets"), which can have a powerful impact for the readers (Racine et al., 2005). This is important because, according to Roskies (2007) neuroimaging studies become increasingly prominent in philosophical and everyday reasoning about behaviour, cognition, and also other (complex) concepts like free will (e.g. Libet, 1999).

Even though there is an increase in use of the results of these studies, it remains questionable whether these results are interpreted correctly. As opposed to what commonly is thought about fMRI results, fMRI does not directly measure neuronal data. Instead it measures the

blood oxygen level dependent (BOLD) signal. This signal comes from hydrogen atoms, which absorb energy at a particular wavelength (Heeger & Ress, 2002). These hydrogen atoms then emit energy at this same frequency until they return to their beginning state, fMRI then measures the sum of this emitted energy (Heeger & Ress, 2002). fMRI does thus not directly measure neuronal activity, but instead measures the timescale of the dephasing water molecules in the brain (Roskies, 2007). See Fig. 2 for a schematic overview of this process.

According to Roskies (2007) non-

January 2018



Fig. 2. Blood oxygen level dependent (BOLD) signal (Heeger & Ress, 2002)

specialists have a commonsense apprehension of neuroimaging data, where fMRI images are thought of as photographs of the brain (and its activity). According to Roskies this association is based on several similarities, as the output of both neuroimaging data and photography (a) result in an image, (b) these images are produced mechanically, and (c) both outputs aim to explain something about the way the world is. However, there are also important differences to be made between photographs and neuroimaging data. Photographs are a direct representation of what has been photographed, the end result is a visual representation of something that was visual to begin with. Brain images on the other hand are a visual representation of brain activity, which does not have any visual properties on its own (Roskies, 2007). In contrast Racine et al., (2005) provides several examples of articles claiming that fMRI results give us visual proof, and that fMRI can (visually) show us what is happening in the brain. Thus, while neuroimages only translate nonvisual information into a visual representation, articles that make use of neuroimage seem to think that neuroimages provide us with visual proof. Furthermore, as explained before, fMRI does not measure brain activity, but instead measures the timescale of the dephasing water molecules in the brain. This means that the measurements made by fMRI need to be translated to brain activity in order to receive brain images that visually represent brain activity. The remaining image is therefore dependent on the beliefs of what the experimenter thinks the neuronal activity of the specific fMRI signals means. So, while most photography is belief-transparent as it is possible to interpret the photo with the information that is contained within the image, neuroimages are belief-opaque, as the information needed to interpret the image is not present within a neuroimage (Roskies, 2007).

That neuroimages are belief-opaque becomes even more clear as Roskies explains that certain techniques used to develop these images are not recoverable from the information inherent in the neuroimage, yet these are all necessary to properly interpret the image. These techniques include (a) *task design:* the tasks involved in the process of generating a neuroimage (which

Master thesis

influence what the data represents), (b) *subtraction:* the functional decomposition of the neuroimage as assumed by the experimenter (neuroimages are always a comparison between multiple datasets), and (c) *statistical analysis:* which include the statistical methods that are employed to get to the end result (e.g. reducing the noise of the fMRI signal). Klein (2010) seems to agree with Roskies, as he states that there are various conceptual problems¹⁴ with neuroimages. Klein (2010) argues that because of these conceptual problems, neuroimages do not provide any evidence necessary that could back up (the often made) conclusions about causality between certain brain regions and cognitive tasks (functional hypotheses). This does not necessary mean that neuroimages are useless, but rather that they should be used to indicate which brain regions need further analysis. According to Klein (2010) it is this further analysis, rather than the neuroimages, that can provide the necessary evidence for functional hypotheses.

From the previous paragraph it becomes clear that, while neuroimages and photographs do have similarities, brain images should not be conceived as photographs of the brain (or brain activity). According to Roskies (2007) it is dangerous that non-specialists see neuroimages on par with photography as it affects how neuroimages are interpreted, with misleading and potentially harmful results. While Roskies does not mention what she means with (non-)specialists, I infer from her article that specialists refer to neuroscientists and the likes, and that therefore (most) philosophers are not included. What furthermore might add to the misinterpreting of neuroimaging results, is the terminology used within some neuroscientific articles, for example the terms "mind-reading" and "brain-reading" (e.g. Norman, Polyn, Detre, & Haxby, 2006; Cox, & Savoy, 2003; Op de Beeck, 2010). The terms "mind-reading" and "brain-reading" are used to describe the process decoding of information from neuroimages (Norman et al., 2006), but may be interpreted quite differently by non-specialists. The danger of misinterpretation does not necessarily exclude neuroscientists, as even neuroscientists do not precisely know what information about brain activity is carried by fMRI signals (e.g. fMRI signals cannot distinguish between excitatory and inhibitory neural activity) (Roskies, 2007). It is therefore wise to question the validity of conclusions and interpretations made by authors (including philosophers) that follow from fMRI and from other brain imaging technologies.

It is important to note that I claim this to be the case for brain imaging technologies in general, and not only for fMRI, as interpretations and conclusions from other brain imaging technologies suffer from the same validity and reliability problems. The claim that we should be concerned about the validity of brain imaging technologies is supported by other authors as well (see Aydin, in press-a), and I will further support this claim in the next section.

Libet experiments

The Libet experiments are probably the best known case where brain imaging technologies are applied to support conclusions made surrounding free will. In this section I will explain what these Libet experiments entail and show that the usage of brain imaging technologies can impact the concept agency. The Libet experiments are relevant to the question of how brain imaging

¹⁴ Some of these conceptual problems are related to the use of null hypothesis significance testing. Master thesis
Page 21 of 38

technologies mediate libertarian theories of free will, as they give a practical example of how brain imaging technologies are used in relation to concepts like free will and agency. Furthermore it will show that the claim made in the previous section (about questioning the validity of conclusions resulting from brain imaging technologies) is not only limited to fMRI but also applies to other brain imaging technologies.

In his experiments Libet used two different brain imaging technologies, namely EEG (electroencephalography) and EMG (magnetoencephalography), to measure the onset time of the so-called readiness potential (Libet, Gleason, Wright, & Pearl, 1983). The readiness potential is the neural activity that becomes visible within the brain just before the participants of his experiments become consciously aware of their decision (their will) to move (Libet et al., 1983). His participants were asked to flex their wrist or fingers and remember the time when they felt the will (the urge) to flex on the position on a oscilloscope clock, while their brain activity was being monitored by a EEG (via electrodes on the participant's scalp). The time of this flexing (which Libet considered to be a simple motor movement), was measured by a EMG (via electrodes on the participants fingers and wrist) and the place on the oscilloscope clock was later translated to a certain amount of time in milliseconds. Libet then found out from the EEG measurements that his participants showed neural activity (which is the onset of the readiness potential) ~550 milliseconds before the onset of the motor act, while the participants only became aware of their decision ~200 milliseconds before the onset of the motor act (Libet et al., 1983; Libet 1999). This means that the readiness potential therefore happens ~350 milliseconds before the participants reported that they had the will to act.

This readiness potential was then used by Libet to claim that these actions already started before the agent could freely will to perform the action, and that therefore the decision to act was not freely done by the agent (O'Connor, 2002). Libet's conclusions are, however, based on certain assumptions (e.g. voluntary acts can only come from conscious decisions), and the wording of the concepts (like readiness-potential) are ambiguous (Mele, 2006)¹⁵. Taking these concerns into account, one could draw various of alternative conclusions from Libet's experiments about the readiness potential (see Balaguer, 2015). My previous made claim that we should question the validity of conclusions and interpretations made by authors following from the results from brain imaging technologies, does therefore not only apply to fMRI but also for other brain imaging technologies.

Later on, Libet (1999) claimed that free will is not necessarily excluded, as an agent becomes aware of their decision (urge) 200 milliseconds before the action, it is still possible for the agent to veto their action in this timeframe. So according to Libet a person can stop the final progress of an action. In other words, human beings may not have free will, but we do have "free won't" (Aydin, in press-a; p.3). Libet made the claim that humans have free won't quite strong, as "[t]he existence of a veto possibility is not in doubt" (p.52). This free won't occurs after the person becomes consciously aware of their action; the person is therefore able to resist the urge.

 ¹⁵ For example, Libet uses words like "urge", "decision", "wish", "intention" and "will" interchangeable to refer to the mental process that people report in his experiment (Mele, 2006). While these words have similarities, they are normally seen as distinct mental processes.
 Master thesis

Therefore, according to Libet's theory agency is limited to our consciousness, and even more specific consciousness that did not originate from unconscious brain activity. However, both the unconscious readiness potential, establishing that we do not have free will, as well as the conscious veto power, establishing that we do have free won't, are based on brain imaging technologies. Not only does this entail a central role for brain activity for free will, it also entails that human agency is dependent on brain activity. In other words, Libet's experiments suggest that human agency can be measured by brain imaging technologies. This means that consciousness, as well as agency are seen as concepts that can be measured by scientific methods. These concepts are therefore not seen as something spiritual or mysterious, but something that has yet to be researched (which is in line with Doyle (2013) as described in the introduction).

The conclusion of Libet that we do not have free will (except for free won't) is based on the assumption that we can only identify¹⁶ ourselves with our consciousness. We can only identify ourselves with our consciousness because, as explained previously, voluntary (willed) acts can only originate from our consciousness. However, the assumption that we only identity ourselves with our unconscious urges and habits (Aydin, in press-b). Take for example a sport (like tennis), the actions a player makes are often unconsciously made. As when one would start thinking of how to handle their racket in tennis, one would suddenly have a harder time to hit the ball. Nevertheless, people would probably disagree with the tennis player if he or she would claim that he or she is not good at tennis, because these movements are made unconsciously, and that only when he or she starts thinking about their movements his or her identity becomes apparent.

We do not only identify ourselves with our (unconscious) skills and achievements, but also with certain character traits that we have. Take a clumsy person. This clumsiness does not come from conscious actions, but happens unconsciously, and to say that someone does not identify himor herself with their clumsiness is often not true. A sudden change in character traits makes this connection between character traits and identity even more clear. If someone suddenly acted clumsily, while he or she normally is not, people (including the person in question) would probably wonder what is wrong with that person, why that person is not acting as usual (and probably wonder if that person might be intoxicated).

It should be noted that it is possible that one does not identity him- or herself with their unconscious behaviour, for example one might not feel that the aggressive thoughts/actions one has are one's own. The same can be said for conscious behaviour as well; for example, someone who consciously cannot resist the urge to eat unhealthily might not identify with these decisions, as he or she might not feel themselves when eating unhealthily (Aydin, in press-b). What therefore seems important for agency is not whether or not the act is initiated consciously or unconsciously, but whether one can identify with one's actions. After all, if an agent is a being (in this case a person) with the capacity to act, but if this person is not identifying with the act, is it really this

¹⁶ Note that identification is important for agency, because (as I will argue a few paragraphs later) an agent is a being with the capacity to act, this being should be able to identify themselves with their actions. If the being cannot identify with the action, is it really this being who is acting? Master thesis Page 23 of 38

person (this being) who is acting? Therefore, if one does not identify with their actions, one might not be performing agency at that time.

So how do these experiments impact agency? According to Aydin (in press-a) opponents as well as advocates of free will draw upon these experiments to support their theories of the existence or non-existence of free will. These experiments suggest that the brains are the locus of agency, they as well emphasise that agency happens only consciously. It is therefore possible that the Libet experiments give rise to a more brain-centred vision towards agency. But the possible impact of these experiments is not only limited within philosophical debates, as articles concerning brain imaging technologies are brought to the attention to multiple disciplines like psychology and physics, but also in media and lay publications (Roskies, 2007). As the Libet experiments can be seen as highly influential, it seems logical to assume that they might have had an influence on our common sense thinking about the world and concepts revolving around free will, which may in the end also influence how we think about concepts like agency.

Conclusions

Within this chapter I showed that it might be wise to question the validity of conclusions and interpretations made by authors, including philosophers, following from results from brain imaging technologies. I also gave a detailed explanation of the Libet experiments, and how the results of this experiment can impact the conception of agency. Furthermore, I used the Libet experiments to show, that the advice to question the validity of conclusions and interpretations made following from the results from brain imaging technologies, does not only account for fMRI but accounts for brain imaging technologies in general. With these pieces I tried to show how brain imaging technologies can affect the way we think about agency and how we might perceive agency. It should now be clear that brain imaging technologies have a certain role within the mediation process. There are however still a few questions left. What exactly does technological mediation entail? How does the effect in perception of agency affect libertarian theories of free will? How exactly do brain imaging technologies mediate libertarian theories of free will? In the next chapter I will try to answer all these questions by explaining what this mediation process exactly entails, and by further exploring how brain imaging technologies impact our concept of agency. I will furthermore propose that brain imaging technologies have strengthened the view that the mind is located in the brain and that this is one of the reasons that there is a current shift within libertarian theories from the agent-causal framework towards the event-causal framework.

5. Technological mediation theory

In the previous chapter I concluded that the Libet experiments might give rise to a more braincentred view towards agency, and that these experiments might also influence our common sense thinking about agency. In this chapter I will make use of the technological mediation theory. This theory entails that technologies (like brain imaging technologies) are not neutral. Normally the influence of these technologies is ignored or pushed to the background, the technical mediation theory allows us to reveal what this influence is. I will start this chapter by explaining the

technological mediation theory in more detail. I will then use the technical mediation theory further develop the claims made in the previous chapter about how brain imaging technologies affect the concept of agency. Furthermore, I will explain how such an effect in perception of agency affects libertarian theories of free will. At last I will explain that the mediation by brain imaging technology has strengthened a recent shift of libertarian theories from agent-causal framework towards the event-causal framework.

Technological mediation

As explained before, the technological mediation theory is a framework that can be used to analyse the various types of relations between human beings, technologies and the world (Verbeek, 2015; Verbeek, 2016). Furthermore, the technological mediation theory presupposes that technologies like fMRI scanners (and other brain imaging technologies), should not be viewed as merely neutral intermediaries (Verbeek, 2015). According to Verbeek (2015) how we perceive and understand certain concepts (e.g. the brain) cannot be understood without taking the mediating role of technologies (in this case BITs) into account. Because I propose (later on in this chapter) that there is a mediation between brain imaging technologies and libertarian theories of free will it is necessary to understanding what the technological mediation theory entails.

In one of his works Verbeek (together with Rosenberger) talks about the mediation of imaging technologies in scientific and medical practice (Rosenberger & Verbeek, 2015). Note that they are not specifically talking about brain imaging technologies, but about (scientific) imaging technologies, and that they therefore also include other technologies than only brain imaging technologies. Nevertheless, Rosenberger and Verbeek (2015) mention that the most expansive set of studies concerning technical mediation is how imaging technologies are used within science and medicine¹⁷. In this work they explain that imaging technologies are used to transform imperceptible aspects of the world into a readable format, namely images. The user of these images receives a transformed experience of this aspect of the world when they try to interpret the image. According to Rosenberger and Verbeek (2015) the user is able to interpret the image in the same way as that one interprets written language; when one is able to interpret the image, the contextual information appears at once and the users "reads" the image as an organised whole. However, the user views a picture that they think adequately represents an aspect of the world, while in actuality they look at a technologically mediated world. Rosenberger and Verbeek (2015) say that scientific images are not simple encounters with the world itself, but claim that these images themselves should rather be understood as technologies that mediate human experience. Furthermore, they say that the creation of such an image is done by an expert; someone that is deeply familiar with how to interpreted these images. The user of this image (when it is not an expert) does not have the knowledge of and familiarity to correctly interpreted the contextual information presented in the image. These findings about the mediation of imaging technologies are in line with the earlier discussion about fMRI in chapter 4.

¹⁷ Rosenberger and Verbeek (2015) provide a list of (examples of) studies concerning imaging technologies on page 32-33 Master thesis

Directly perceiving agency

Brain imaging technologies give us the feeling that we can have a direct view of what is happening within a brain. The brain consists of many neurons, which are cells that can communicate (transfer information) with other neurons via electronic impulses (within the neuron) and chemical signals (from one neuron to another neuron via a synapse), all the neurons create a system, which is called the nervous system. Within this nervous system lies not only the brain, but also the sensory organs as well as the spinal cord, which is an essential part in our ability to move. As all nerves are connected to the brain, and the brain can send and receive information via these nerves the brain could therefore be seen as the centre, or operating organ, of the nervous system. As both sensory functions and bodily functions are connected to the brain, it could be stated that the brain gives an agent the capacity to act, as without the brain, the agent could not make bodily movements nor gain information from our sensory organs¹⁸. Note that this does not necessary mean that agency is limited to the brain.

Because brain imaging technologies give us the feeling that we can directly see what is happening within the brain, and because the brain gives an agent the capacity to act, we get the idea that we can also directly perceive agency. However, as explained in the section fMRI (chapter 4), we do not directly perceive what is happening within the brain, and therefore we are unable to directly perceive agency with these brain imaging technologies. It is probably safe to assume that the conclusions of Roskies (2007) and Klein (2010) (and similar articles) are not widespread known, especially among laypeople, but also possibly among philosophers (see also Wardlaw et al., 2011). Many people are therefore not aware that we cannot directly perceive agency, and even if they are aware they are already influenced by the idea that we (in principle) can directly perceive agency. One might even claim that it will only take some time until new technologies are developed that are able to let us directly perceive what is happening within the brain (e.g. Doyle, 2013). Therefore, the mere existence of brain imaging technologies helps to shape how agency is perceived and how this agency is interpreted by humans.

Limiting agency to the brain

Aydin (in press-a) argues about how brain imaging technologies mediate a more general conception of free will. Within his article he tries to display the questionability of the current framework used in free will debates and to propose an alternative framework that recognises the mediation of brain imaging technologies. Aydin explains how brain imaging technologies mediate our conception of the brain in free will debates. According to Aydin, brain-images are conceived as a direct representation of a brain unaffected by external influences (corresponding with the arguments that I make in the previous chapter). He continues by stating that this leads to the idea that not consciousness but the brain is often seen as the locus of the self. Aydin proposes that we should not interpreted free will in terms of self-determination but rather in terms of self-formation (more thoroughly discussed in Aydin, in press-b).

 ¹⁸ To not complicate matters, I will for this argument ignore the possibility of actions that do not necessarily require our sensory organs or our ability to move.
 Master thesis
 Page 26 of 38

The notion of self-formation is one that I have yet to explain. I should note that it is not my aim to provide an exhaustive discussion about the self and self-formation. The notions of self and self-formation in this thesis is used in order to support my claim further on in this section that agency is being limited to the brain. For the sake of this thesis I will now provide a basic framework of what self-formation is and how it is used within libertarian theories. As mentioned in chapter 2 libertarian Robert Kane uses the term self-formation as a central concept to explain free will. According to Kane (2005) (the freedom of) self-formation follows logically from (the freedom of) self-determination. Kane (2005) defines (the freedom of) self-determination as: "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." (p. 172). In other words, self-determination is the ability to act freely on the basis on your own formed will. However, self-determination does not require the power to do otherwise at the moment of a decision (Kane, 2005). As it is possible that the will of an agent has formed in such a way that the agent cannot do otherwise (Kane, 2005). This is a problem for Kane (2005) because libertarian free will requires an agent to be able to do otherwise (to freely choose an alternative path). He therefore arrives at a definition of (the freedom of) self-formation, which is: "the power to form one's own will [(character, motives, and purposes)] 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" (p.172). In other words, self-formation is the formation of our wills; how we became the kinds of persons we are, with the wills we now have (Kane, n.d.).

Now to return to the discussion at hand. Aydin (in press-a) proposes that we should not interpret free will in terms of self-determination but rather in terms of self-formation. According to Aydin one step in this process is to acknowledge that "brain scanners in themselves are not sufficient to display a brain that can be seen as the locus of the self [...] but rather contribute to the formation of a self and the advancement of its freedom" (p. 2). This thought is in line with my conclusions in the section "fMRI", where I state that brain images should not be seen as a direct representation of what is happening within the brain, and that it is therefore wise to question the validity of conclusions and interpretations that follow from the use of brain imaging technologies.

As Aydin's claims are about free will in general, he therefore uses the concept of self, which can be seen as pivotal to free will. I claim however that a similar argument can be made at a more specific level, namely about the concept of agency: Brain imaging technologies are not neutral technologies that display agency, but instead contribute to the formation of agency and therefore influence our perception of agency. What I mean with this contribution to the formation of agency, is: that brain imaging technologies change how we perceive and understand the concept agency in such a way, that brain imaging technologies themselves affect our perception of agency. In other words, there is a mediation between brain imaging technologies and agency, and because the concept of agency is pivotal to libertarian theories of free will (see chapter 3 "Agency and libertarian theories of free will"), this leads to the mediation between brain imaging technologies and libertarian theories of free will.

This mediation between brain imaging technologies and agency is possible because as explained in chapter 4 (Brain imaging technologies and agency) brain imaging technologies change how we perceive brains. One of these changes, as suggested by Aydin (in press-a), is that the brain is more often seen as the locus of the self. This can, among others, be seen in news articles. Racine et al. (2005) performed a meta-analysis, in which he noticed that a lot of news articles report about fMRI research where concepts like identity and self are reduced to the brain. Among recent Libertarian theories the brain is often given a more pivotal role, and is often seen as the locus of the mind. Examples of contemporary libertarian accounts where the brain is given a more central role are given in the introduction of this thesis, and include authors like Robert Kane, Robert Doyle and Martin Heisenberg. Kane (2013) states that processes that constitute to introspection take place in the brain. And introspection is "a means of learning about one's own currently ongoing, or perhaps very recently past, mental states or processes" (Schwitzgebel, 2016, p. 1). More importantly, introspection is one of the core functions that we often attribute to the mind. I would like to note that especially Robert Kane is important if we consider the state of libertarian accounts of free will, as Robert Kane is acknowledged as the leading spokesman for Libertarianism (Doyle, 2011e). Therefore, if Robert Kane gives the brain a more central role within libertarian theories, and proposes that the brain is the locus of functions of the mind, it will most likely lead to other libertarians doing the same.

Within these theories the processes that the mind performs are either located in the brain, or can be traced back to brain processes, limiting the mind to the brain. Furthermore, the self, as well as the mind, are two concepts that are strongly linked to the concept of agency. The capacity for an agent to act is (often) dependent on the agent's mental states (see chapter 3), and it is the actions of the agent that are often said to form the self (e.g. Kane, 2007 as explained in the section "Indeterminism" in chapter 2). Limiting the self or the mind to the brain does therefore also limit agency to the brain.

I claim that brain imaging technologies have strengthened the view that the mind is located in the brain. As we hear more news involving brain imaging technologies, read about research done via brain imaging technologies etc., different concepts of the mind seem to collapse. In the end technologies like brain imaging technologies can change how we use language, while the technologies themselves remain in the background. Because they remain in the background we do not notice how this affects our thoughts and interpretations and conclusions we make. An example of how this affects our conclusions is visible if we look at the end of the section of two-stage models in chapter 3. As stated in that section Doyle claims that new possibilities exist only within our *minds*, and that neuroscientists might eventually be able to predict our thoughts and decisions if they could resolve the finest details of information storage in our *brains*. Thereby Doyle does not only limit the mind to the brain, but also states that it is in principle possible to use neuroscience (e.g. by brain imaging technologies) to explain our minds, and to read our thoughts.

Shift towards event-causal libertarian theories

The change in how agency is perceived, via a change in how our brains are perceived, as explained in the previous section, might seem subtle. In this section I will show that even if it might seem subtle, it contributes greatly to the mediation of libertarian theories of free will. In this section I propose that this mediation between brain imaging technologies and agency has strengthened a recent shift of libertarian theories from the agent-causal framework towards the event-causal framework.

In chapter 3 (Agency and libertarian theories of free will) I noted that the agent-causal framework is more popular among contemporary libertarian account of free will. This might seem in conflict with my proposal that there is a recent shift visible towards the event-causal framework. I should however note that both statements can be true at the same time. That there is a recent shift does not necessarily mean that there are more event-causal libertarian theories than there are agent-causal libertarian theories, it merely means that there is an increase visible in the amount of event-causal libertarian theories, which might eventually make the event-causal framework more popular among libertarians. Furthermore, as I will explain next, some authors might think that they are using the agent-causal framework, while in actuality they are using the event-causal framework.

Because the origin of this shift, the mediation of brain imaging technologies, is pushed to the background, the shift itself might not be noticed by the libertarian authors themselves. Instead it might cause a change in what authors think the agent-causal framework entails. This becomes evident if we look the two-stage models described in chapter 3 of this thesis. In that chapter I made the claim that these models make use of the event-causal framework. However, the main proponent of these two-stage models, namely Robert "Bob" Doyle, claims that two-stage models¹⁹ is "perhaps less 'event causal' and more 'agent-causal,' because the agent has creative powers during the extended 'moment of choice'" (Doyle, 2013, p. 251). In this quote "moment of choice" refers to the second stage of the two-stage model, while creative powers probably refer to the creative process described in Doyle's definition of free will in the previous paragraph. I would think the reason Doyle puts two-stage models in the agent-causal framework is because of the first section of his definition of free will. In this part he states that humans have the power to generate alternative possibilities, and therefore one could argue that the chosen decisions originate from the agent. However, it is from the last part of his definition that it becomes quite apparent that twostage models (or at least his two-stage model) make use of the event-causal framework instead of the agent-causal framework. In this part he states that the possibilities are evaluated, after which one is "willed", by which he describes "willed" as selected or chosen by reasons, motives or desires. One of the main distinctions between the agent-causal framework and the event-causal framework is that the agent-causal framework can be seen as non-reductionist, while within the event-causal framework the agent gets reduced to mental states or events. It is clear that within his definition of

¹⁹ It might be possible that with "two-stage models" he was merely referring to his own two-stage model. Master thesis

free will Doyle reduced the agent's "will" to mental states and events, as they are adequately determined (in other words caused) by reasons, motives or desires.

As explained before the agent-causal framework provides a non-reductive account of agency, it is the agent (as a substance) and not something else that is the cause of its own actions. It seems possible to see the mind as the location of where agency is performed in the agent-causal framework, as long as this means that the agent is equal to the mind, or somehow gives orders to the mind, as this does not reduce the causal powers of the agent. It is however quite difficult to incorporate the brain in the same way, as the brain can be reduced towards brain activity, impulses, thoughts etc. some of which can originate from outside the body.

Conclusions

Throughout this text I explained that more often the brain is given a pivotal role within libertarian theories, and that the brain is more often seen as the locus of the mind and the self. In the section "Limiting agency to the brain" I concluded that brain imaging technologies strengthen this change. And in the same section I explained that seeing the brain as the locus of the mind and the self also entails that the brain is the locus of agency. From the previous paragraph it should also be clear that if the brain is the locus of agency than it is not possible for a theory that endorses this idea to be an agent-causal theory. Libertarian accounts that use the brain as either the locus of the self, or as locus of the mind make therefore use of the event-causal framework. I therefore propose that the mediation between brain imaging technologies and agency has strengthened a recent shift of libertarian theories from the agent-causal framework towards the event-causal framework. Furthermore, I claim that this shift might not be noticed in such, as the origin of the shift (the brain imaging technologies) are pushed towards the background. This causes the authors to either mistakenly think that their theories do fit within the agent-causal framework, or for authors to change the meaning of agent-causality. This might be the most noticeable among libertarian accounts that have a relation to or make conclusions inferred from the measurements of brain imaging technologies.

6. Conclusion

Throughout this thesis I have tried to give an answer to the question how brain imaging technologies mediate libertarian accounts of free will. In order to answer this question I proposed three subquestions: (a) In what way is the concept agency central to libertarian theories of free will?; (b) How do brain imaging technologies affect the way we think about agency?; and (c) How does such an effect in the perception of the concept agency affect libertarian theories of free will?

Before I could go to the first subquestion I needed to provide a framework of libertarianism. To do so I explained that there are three other categories in which free will theories are often situated, namely: a) Hard determinism; b) Compatibilism; and c) Hard incompatibilism. I described the four categories and how libertarianism differs from these categories. In short, libertarians are incompatibilists that believe free will exists, requiring indeterminism of some sort. Indeterminism entails that the future is not set is stone, and that (at least some) events are not necessarily bound by the laws of nature, which allows persons/agents to choose from alternative paths in the future. I explained that indeterminism does not rule out causation nor does indeterministic causation only entail luck or chance. In short, I explained how libertarianism differs from other positions, and how the relevant terminology should (and should not) be interpreted and used.

To answer the first subquestion I needed to establish that the concept of agency is central to libertarian theories. In order to do so I explained that agency is the manifestation of the capacity of a being to act, and furthermore explained how the standard *conception* and the standard *theory* of action can be translated towards a standard *conception* and standard *theory* of agency. I established the link with agency and libertarianism by explaining that there are three frameworks of agency, which are identically named as the three different kinds of libertarian theories, namely: (a) agent-causal; (b) noncausal (volitionist); and (c) event-causal. The agent-causal framework is the most popular among contemporary libertarian theories, while the event-causal framework is the most popular among contemporary philosophy in general. The agent-causal framework provides a non-reductive account of agency; proponents argue that while reasons and intentions might influence how an agent acts, the source of the agency lies in the power to initiate. This causation by the agent is seen as causation by a substance; the agent is therefore the *causa sui* of its own decisions. According to the event-causal framework, actions can always be reduced to certain events, and events are only actions when there is a causal relation between the event and the mental states of the agent. Libertarian event-causal theories require that at least some of the events that cause actions are indeterministically caused, which gives the agent the possibility to do otherwise. In short, the answer to the first subquestion is that agency is central to libertarian theories because both libertarian theories and agency can be categorised by the same metaphysical frameworks.

To answer the second subquestion I needed to show how brain imaging technologies affect the way we think about and perceive agency. In order to do so I went into detail about fMRI, which is seen as a non-invasive method to measure neuronal activity in the human brain. The results of fMRI, which are non-visual neural information, are presented in an easy accessible visual way, which is a problem for lay people and others who might lack the necessary expertise to correctly interpret the images. Furthermore, fMRI does not, as commonly thought, measure neuronal data, but instead measures the timescale of the dephasing water molecules in the brain. I furthermore explained that non-specialists (including philosophers) have a commonsense apprehension of neuroimaging data, where fMRI images are seen as photographs of the brain. However, this analogical way of seeing fMRI images is incorrect, as photographs are belief-transparent while neuroimages are belief-opaque. It is therefore wise to question the validity of conclusions and interpretations made by authors (including philosophers) that follow from fMRI as well as for other brain imaging technologies.

I furthermore went into detail about the Libet experiments. Libet used two different brain imaging technologies (EEG and EMG) to measure the onset time of the readiness potential, which is the neural activity that becomes visible within the brain of an agent just before their awareness to

move. This lead Libet to claim that motor movements are iniated before the agent could freely will to perform these actions. Libet later claimed that free will might not be necessary excluded as an agent still has the possibility to stop the final process of an (not freely initiated) action. However, because Libet's conclusions are based on certain assumptions, and ambiguous worded concepts, one could draw various alternative conclusions from Libet's experiments. As I explained, according to Libet's theory agency is limited to our consciousness, specifically to consciousness that did not originate from unconscious brain activity. Furthermore, because Libet bases all of his conclusions on results on brain imaging technologies, Libet's experiments suggest that human agency can be measured by brain imaging technologies. Therefore, these experiments suggest that the brains are the locus of agency, as well as that agency happens only consciously. Because both opponents and advocates of free will draw upon these experiments it is possible that the Libet experiments give rise to more brain-centred vision towards agency. In short, the answer to the second subquestion is that research done with brain imaging technologies can influence how we think and perceive agency, as they could for example lead to a more brain-centred view towards agency, and influence our common sense thinking about agency. We should however be careful and question the validity of conclusions and interpretations made by authors that follow from brain imaging technologies.

In order to answer the last subquestion I first needed to provide a framework about the technological mediation theory. I explained that the technological mediation theory entails that (brain imaging) technologies are not as neutral as we like to think they are, and that it allows us to reveal the influence of these technologies, which is often ignored or pushes to the background. Furthermore, how we perceive and understand concepts like the brain and agency cannot be understood without taking the mediation role of technologies into account. I noted that images derived from imaging technologies should not be understood as simple encounters of the world itself, but as technologies that mediate human experience. I furthermore explained that brain imaging technologies give us the feeling that we can directly perceive the brain, and therefore also agency. However as concluded earlier, brain imaging technologies do not provide us with a direct view of the brain, but rather translate certain non-visual data into images. Many people are however not aware that we cannot directly perceive agency, and even if they do they might already be influenced by the idea of directly perceiving agency. Therefore, the fact that brain imaging technologies exist, does already shape our perception and interpretation of agency.

Furthermore, I explained that authors that do rely on results from brain imaging technologies do not take into account that brain imaging technologies are not neutral tools to measure properties of humans and the world around us, but co-determine what they measure, and therefore frame concepts like the brain in a certain way. I make the claim that brain imaging technologies are not neutral technologies that display agency, but instead contribute to the formation of agency and therefore influence our perception of agency. And that this leads to the mediation between brain imaging technologies and libertarian theories of free will. Brain imaging technologies change how we perceive brains in various ways. Brains are given a more pivotal role, and the brain is often seen as the locus of the self or the locus of the mind. The idea that Robert Kane (leading spokesman for Libertarianism) also gives the brain a more central role and sees the

brain as the locus of the mind most likely means that other libertarians will do the same. I point out that these authors limit the self or the mind to the brain. Because the self and the mind are two concepts strongly linked to the concept of agency, limiting the self or the mind to the brain, also limits agency to the brain. I furthermore claim that it is most likely because of brain imaging technologies that the mind is now often thought to be located in the brain, but that it is questionable whether the brain is all there is when we talk about the mind. The answer to the last subquestion is shortly after introduced, as I claim that the more brain-centred view, as well as seeing the brain as the locus of the mind strengthens a shift in libertarian theories from the agent-causal framework to the event-causal framework.

So now to answer the main question: "How do brain imaging technologies mediate libertarian accounts of free will?". Brain imaging technologies strengthen a change in contemporary libertarian theories. This change is that the brain is given a more pivotal role within libertarian theories, leading to seeing the brain as the locus of the mind and the self. Seeing the brain as the locus of the mind and the self also entails that the brain is seen as the locus of agency. Libertarian theories that endorse the idea that the brain is the locus of agency cannot fit into the agent-causal framework, and therefore make use of the event-causal framework. I claim therefore that this mediation between brain imaging technologies and agency strengthens a recent shift of libertarian theories from the agent-causal framework towards the event-causal framework. This shift itself might not even be noticed by the libertarian authors themselves, as the origin of the shift, the brain imaging technologies, is pushed to the background.

Discussion and follow-up research

There are a few discussion points that I should address about this thesis. Firstly, the information about the two-stage models that I describe in my thesis originates mostly from Doyle. Furthermore, while Doyle provides a whole list of authors that describe a theory that involves a two-stage model, these authors do not often state in their articles that their theory is a two-stage model of free will. Secondly, I claim that the concept of agency and as well as libertarianism can be categorised within three metaphysical frameworks: The event-causal framework, the agent-causal framework, and the noncausal framework. It is plausible that there are some libertarian theories that do not fit within these categories, or that there is some form of agency that is different from the ones described in this thesis. Thirdly I assume that because the spokesman of libertarianism (Robert Kane) gives the brain a more central role within libertarian theories, and proposes that the brain is the locus of the mind, it will most likely lead to other libertarians doing the same. This assumption might not be true, but it would be hard to either prove or disprove this assumption.

Another point of discussion is that I generalise the claim "that it is wise to question the validity of conclusions and interpretations made by authors that follow from fMRI" to all brain imaging technologies. While I support this generalisation in the section about the Libet experiments, I do not go into depth about all the other existing brain imaging technologies, like I did with fMRI. It might be possible that there are some brain imaging technologies for which the validity of the derived conclusions and interpretations do not have to be questioned in the same

way. Further research could be done in order to validate my claim that it is wise question the validity of conclusions and interpretations made by authors that follow from all the other brain imaging technologies. Furthermore, it could be worthwhile to compare the brain imaging technologies and to see whether there are important differences concerning the validity of these brain imaging technologies and to check whether they may be differences in how these technologies mediate the concept of agency.

Next to the previous proposed follow-ups, my main proposal for a follow-up research is as follows. In order to more reliable assess whether or not there is a recent shift from agent-causal libertarianism towards event-causal libertarianism an in-depth comparison of all libertarian articles/theories is necessary. In this in-depth comparison it should be investigated whether these authors make use of an agent-causal framework or an event-causal framework as described in this thesis. Furthermore, it might be worthwhile to see whether they specify themselves to which framework they endorse, and to see if their thoughts about their endorsed framework complies within the description in this thesis. One of the difficulties of this research is that it is not easy to search for libertarian articles, as authors use different names to refer to libertarianism. Furthermore, the term libertarianism is currently more used in the political sense, which might make the search for articles even harder.

Bibliography

- Aydin, C. (in press-a). From camera obscura to fMRI: How brain imaging technologies mediate 'free will'.
- Aydin, C. (in press-b). Freedom as self-formation: Reframing 'free will' from the perspective of the mediated self.
- Balaguer, M. (2004). A Coherent, Naturalistic, and Plausible Formulation of Libertarian Free Will. *Nous*, *38*(3), 379-406. doi:10.1111/j.0029-4624.2004.00475.x
- Balaguer, M. (2015). Determinism and Its Relevance to the Free-Will Question. *Handbook of Neuroethics*, 231-251. doi:10.1007/978-94-007-4707-4_121
- Canli, T., Zhao, Z., Desmond, J. E., Kang, E., Gross, J., & Gabrieli, J. D. E. (2001). An fMRI study of personality influences on brain reactivity to emotional stimuli. *Behavioral Neuroscience*, *115*(1), 33-42. doi:10.1037/0735-7044.115.1.33
- Clarke, R., & Capes, J. (2017). Incompatibilist (Nondeterministic) Theories of Free Will. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition ed.). Retrieved from https://plato.stanford.edu/archives/spr2017/entries/incompatibilism-theories/.
- Cox, D. D., & Savoy, R. L. (2003). Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. *NeuroImage*, 19(2), 261-270. doi:10.1016/s1053-8119(03)00049-1
- Doyle, R. O. (n.d.-a). Mind-Body Identity Theory. *The Information Philospher*. Retrieved Oktober 23, 2017, from http://www.informationphilosopher.com/knowledge/identity_theory.html
- Doyle, R. O. (n.d.-b). Two-Stage Models for Free Will. *The Information Philospher*. Retrieved February 20, 2017, from http://www.informationphilosopher.com/freedom/twostage_models.html
- Doyle, R. O. (2011a). Libertarianism. *Free Will: the scandal in philosophy*. (pp. 152-155). Cambridge, MA: The Information Press.
- Doyle, R. O. (2011b). A Taxonomy of Free Will Positions. *Free Will: the scandal in philosophy*. (pp. 58-67). Cambridge, MA: The Information Press.
- Doyle, R. O. (2011c). Freedom. *Free Will: the scandal in philosophy*. (pp. 18-25). Cambridge, MA: The Information Press.
- Doyle, R. O. (2011d). Two-Stage Models of Free Will. *Free Will: the scandal in philosophy*. (pp. 160-185). Cambridge, MA: The Information Press.
- Doyle, R. O. (2011e). History of the Free Will Problem. *Free Will: the scandal in philosophy.* (pp. 68-131). Cambridge, MA: The Information Press.
- Doyle, R. O. (2013). The Two-Stage Model to the Problem of Free Will: How Behavioral Freedom in Lower Animals Has Evolved to Become Free Will in Humans and Higher Animals. In A. Suarez & P. Adams (Eds.), *Is science compatible with free will?: Exploring free will and consciousness in the light of quantum physics and neuroscience* (pp. 235-254): Springer Science & Business Media. doi:10.1007/978-1-4614-5212-6_16
- Fischer, J. M., Kane, R., Pereboom, D., & Vargas, M. (2007). A Brief Introduction to Some Terms and Concepts. *Four views on free will*. Malden, MA: Blackwell Publishing.

- Gallagher, S. (2007). The Natural Philosophy of Agency. *Philosophy Compass, 2*(2), 347-357. doi: 10.1111/j.1747-9991.2007.00067.x
- Gordon, E. (1999). Brain imaging technologies: how, what, when and why? *Australian & New Zealand Journal of Psychiatry*, *33*(2), 187-196. doi:10.1046/j.1440-1614.1999.00557.x
- Heeger, D. J., & Ress, D. (2002). What does fMRI tell us about neuronal activity? *Nature Reviews Neuroscience*, *3*(2), 142-151. doi:10.1038/nrn730
- Heisenberg, M. (2009). Is free will an illusion? Nature, 459(7244), 164-165. doi:10.1038/459164a
- Heisenberg, M. (2013). The Origin of Freedom in Animal Behaviour. In A. Suarez & P. Adams (Eds.), *Is science compatible with free will?: Exploring free will and consciousness in the light of quantum physics and neuroscience* (pp. 95-103): Springer Science & Business Media. doi:10.1007/978-1-4614-5212-6_7
- Hoefer, C. (2016). Causal Determinism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition ed.). Retrieved from http://plato.stanford.edu/archives/spr2016/entries/determinism-causal.
- Ihde, D. (1990). *Technology and the lifeworld: From garden to earth*. Bloomington, IN: Indiana University Press.
- James, W. (1897). The dilemma of determinism. *The will to believe and other essays in popular philosophy* (pp. 145-183). Salt Lake City, UT: Project Gutenberg.
- Kane, R. (n.d.). Free will: an achievement over indeterminacy. Slate: Are we free? Essays & Opinions. Retrieved from http://www.slate.com/bigideas/are-we-free/essays-andopinions/robert-kane-opinion
- Kane, R. (2005). A contemporary introduction to free will. NY: Oxford: Oxford University Press.
- Kane, R. (2007). Libertarianism. In J. M. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), Four views on free will (pp. 85-125). Malden, MA: Blackwell Publishing.
- Kane, R. (2013). Can a Traditional Libertarian or Incompatibilist Free Will Be Reconciled with Modern Science? Steps Toward a Positive Answer. In A. Suarez & P. Adams (Eds.), *Is* science compatible with free will?: Exploring free will and consciousness in the light of quantum physics and neuroscience (pp. 255-272): Springer Science & Business Media. doi: 10.1007/978-1-4614-5212-6_17
- Klein, C. (2010). Images Are Not the Evidence in Neuroimaging. *The British Journal for the Philosophy of Science*, *61*(2), 265-278. doi:10.1093/bjps/axp035
- Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral activity (readiness-potential). The unconscious initiation of a freely voluntary act. *Brain*, *106*(3), 623-642. doi:10.1093/brain/106.3.623
- Libet, B. (1999). Do we have free will? Journal of consciousness studies, 6(8-9), 47-57.
- Markosian, N. (1999). A compatibilist version of the theory of agent causation. *Pacific Philosophical Quarterly*, *80*(3), 257-277. doi:10.1111/1468-0114.00083
- McKenna, M. (2004). Compatibilism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition ed.). Retrieved from http://plato.stanford.edu/archives/ win2016/entries/compatibilism/.

- Mele, A. R. (2003). Introduction. *Motivation and agency* (pp. 3-7). New York NY: Oxford University Press.
- Mele, A. R. (2006). Free will: Theories, analysis, and data. In S. Pockett, W. P. Banks, & S.
 Gallagher (Eds.), *Does consciousness cause behavior* (pp. 187-205). Cambridge MA: MIT Press.
- Nelkin, D. K. (2013). Moral Luck. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2013 Edition ed.). Retrieved from https://plato.stanford.edu/archives/win2013/ entries/moral-luck/.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in cognitive sciences*, *10*(9), 424-430. doi:10.1016/j.tics.2006.07.005
- O'Connor, T. (2000). *Persons and causes: The metaphysics of free will*. Oxford University Press on Demand.
- O'Connor, T. (2016). Causal Determinism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition ed.). Retrieved from https://plato.stanford.edu/archives/sum2016/entries/freewill.
- Op de Beeck, H. P. (2010). Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *NeuroImage*, *49*(3), 1943-1948. doi:10.1016/j.neuroimage. 2009.02.047
- Pardo, M. S. & Patterson, D. (2013). Minds, brains, and law: The conceptual foundations of law and neuroscience. Oxford, UK: Oxford University Press. doi: 10.1093/acprof:oso/ 9780199812134.001.0001
- Pauen, M. (2008). Self-Determination. Free Will, Responsibility, and Determinism. *Synthesis philosophica*, *22*(2), 455-475.
- Pereboom, D. (2007). Hard Incompatibilism. In J. M. Fischer, R. Kane, D. Pereboom, & M. Vargas (Eds.), *Four views on free will* (pp. 85-125). Malden, MA: Blackwell Publishing.
- Pérez Otero, M. (2016). How Indeterminism Could Help Incompatibilism on Free Action. *Dialectica*, *70*(2), 169-184. doi:10.1111/1746-8361.12142
- Popper, K. R. (1950). Indeterminism in quantum physics and in classical physics. Part I. *The British Journal for the Philosophy of Science*, *1*(2), 117-133.
- Racine, E., Bar-Ilan, O., & Illes, J. (2005). fMRI in the public eye. *Nat Rev Neurosci*, *6*(2), 159-164. doi:10.1038/nrn1609
- Rosenberger, R., & Verbeek, P.-P. (2015). A field guide to postphenomenology. *Postphenomenological investigations: Essays on human-technology relations*, 9-42.
- Roskies, A. L. (2007). Are Neuroimages Like Photographs of the Brain? *Philosophy of Science*, 74(5), 860-872. doi:10.1086/525627
- Schlosser, M. (2015). Agency. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition ed.). Retrieved from https://plato.stanford.edu/archives/fall2015/entries/agency

- Smilansky, S. (2001). IV^{*}—Free Will: From Nature to Illusion. *Proceedings of the Aristotelian Society, 101*(1), 71-95. doi:10.1111/j.0066-7372.2003.00022.x
- Schwitzgebel, E. (2016). Introspection. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition ed.). Retrieved from https://plato.stanford.edu/archives/ win2016/entries/introspection/
- Timpe, K. (n.d.). Free Will. *The Internet Encyclopedia of Philosophy*. Retrieved December 18, 2016, from http://www.iep.utm.edu/freewill/
- Verbeek, P.-P. (2005). A material hermeneutic. What things do: Philosophical reflections on technology, agency, and design (pp. 121-146). Pennsylvania: Pennsylvania State University Press.
- Verbeek, P.-P. (2015). COVER STORY: Beyond interaction: a short introduction to mediation theory. *Interactions*, *22*(3), 26-31. doi:10.1145/2751314
- Verbeek, P.-P. (2016). *Mediation theory*. Retrieved April 12, 2017, from https:// ppverbeek.wordpress.com/mediation-theory/
- Vihvelin, K. (2015). Arguments for Incompatibilism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition ed.). Retrieved from https://plato.stanford.edu/archives/fall2015/entries/incompatibilism-arguments/.
- Wardlaw, J. M., O'Connell, G., Shuler, K., DeWilde, J., Haley, J., Escobar, O., . . . Schafer, B. (2011).
 "Can it read my mind?" What do the public and experts think of the current (mis)uses of neuroimaging? *PLoS One, 6*(10), e25829. doi:10.1371/journal.pone.0025829
- Whittle, A. N. N. (2016). A Defense of Substance Causation. *Journal of the American Philosophical Association*, *2*(1), 1-20. doi:10.1017/apa.2016.1