Can Agents be Causes? A Critique of Timothy O’Connor’s Theory of Free Action.

By Auke Alesander Montesano Montessori

Study Program: Honours Philosophy - Bachelor (University Utrecht)
Student number: 4300939
Supervisor: Niels van Miltenburg
ECTS: 15
Date: 24 August 2017

Table of Contents

Introduction.. 4

I O'Connor's Theory of Agent-causation. 6

II Coherence.. 8

III Explaining Freedom of Action 12

IV Explaining the Relation between Reasons and Action... 16

V Empirical Data and Agent-Causation 22

Conclusion 25

Literature.. 27

Abstract

In this thesis, I critically asses Timothy O'Connor's theory of free action. O'Connor argues that agents have the power to cause action-triggering intentions, which in turn cause the occurrence of action. The agent is free because she controls the action through her agent-causal powers. I discuss the coherence of O'Connor's theory, his account of freedom and reasons for actions and whether his theory fits with empirical data. I conclude that the theory is coherent and intelligible, but that some of its explanations concerning human action are not satisfactory. O'Connor's account of the relation between reasons and action is the weakest part of his overall theory. Where some of his other explanations were merely unsatisfactory, his theory of reasons and actions is, as it stands, completely inadequate. Overall, it would be a mistake to dismiss O'Connor's theory lightly since it generally is a coherent theory that can give us an explanation concerning free will. But, given the fact that a number of O'Connor's explanations are unsatisfactory, I suggest looking for alternative theories that can provide us with better results.

Introduction

How does human action work? Do we have free will and if so, what does free will consist in? What exactly is the link between action and reasons for action? This thesis deals with one specific subspecies of theories that try to answer these questions: agent-causal theories.
Some philosophers, most notably Donald Davidson[1], have thought that human action is caused by the reasons or desires of the agent. Agent-causalists deny this. According to them, human action is caused by the agent herself, although not necessarily directly, and not only by her reasons or desires. These actions are free because the agent herself was not caused to act by some event. She was the origin of the action and she has control over it. Thus, agent-causalists argue, she is free.[2]
Although agent-causalists generally agree on the above picture, several subtly distinct forms of agent-causal theories have been created over the course of time. For reasons of space, clarity and convenience, this thesis will focus on the agent-causal theory of Timothy O'Connor. He is one of the most important contemporary defenders of agent-causal theories of free will.[3] Roughly, he contends that agents have the power to cause an action-triggering intention, which in turn causes the action to occur. Reasons and desires do not play any direct causal role in the coming about of actions, though they can influence the agent in her decision.[4]
The aim of this thesis is to critically assess O'Connor's theory. What are its philosophical merits and what are its pitfalls? Is it coherent? Does it fit current empirical findings? These are the type of questions I shall answer here.
I will open this thesis with a short sketch of O'Connor's theory, just to make clear what exactly we are talking about.
Next, in part II, I investigate whether the stated theory is actually coherent. There are some philosophers who consider agent-causation to be incoherent. I will argue that these philosophers are mistaken and that O'Connor's theory is coherent. However, some of the complaints point us to the fact that O'Connor cannot explain all facets of action equally well.
Part III and IV will critically assess O'Connor's account of the freedom of action and the relation between reason and action. Can he actually show how we can be free and reasonable while acting? I claim that O'Connor can intelligibly explain our freedom. This despite claims that this freedom is indistinguishable from mere chance. The same cannot be said about his account of the relation between reason and action. This account, as can be shown by a number of examples, is clearly not adequate.
Part V will deal with the empirical data we have about ourselves and about how we perform actions. I briefly describe the relation between the current status of the empirical sciences and agent-causal theories. But I mainly focus on phenomenology, because O'Connor thinks that our experience of acting speaks in favour of agent-causal theories of action. I propose that more systematic research on our experience of acting should be done to see whether O'Connor's optimism is justified.
The last part is the conclusion, in which I look back on the virtues and vices of O'Connor's theory. I conclude that O'Connor's theory cannot be described as a failure. He succeeds in giving a coherent and intelligible agent-causal theory on free human action. This is a significant accomplishment. Still, his explanations are not always as satisfactory as we would like to see and his theory on the relation between reasons and actions is inadequate. These problems, added together, are serious enough to warrant the suggestion to look elsewhere for a better theory on action. Whether alternative theories

---

[1] Donald Davidson, ''Actions, Reasons and Causes,'' *The Journal of Philosophy* 60 (1963), 685-700.
[2] Niels Van Miltenburg, *Freedom in Action,* (Utrecht: Quaestiones Infinitae, 2015), 147-148.
[3] Another example of a prominent agent-causal theorist is Randolph Clarke, though he is not fully convinced by the virtue of the concept of agent-causation. Randolph Clarke, *Libertarian Accounts of Free Will*, (Oxford: Oxford University Press, 2003).
[4] Timothy O'Connor, *Persons and Causes,* (Oxford: Oxford University Press, 2000).

fare any better than O’Connor is hard to say, at least within the boundaries of this thesis. Still, at least we shall know the merits of O’Connor’s theory in detail.

I O’Connor’s Theory of Agent-causation.

The basic idea of agent-causation is fairly straightforward. The agent has the ontologically primitive causal capability or power to create an action triggering intention. This intention goes on, in an immediate fashion, to cause the action. Thus, through this action-triggering intention, the agent is in control of the action and is free.[5] O’Connor formulates this conception of agent-causation in the following manner:

‘’O’Connor contends that agent causation (whether actual or merely possible) is an ontologically primitive type of causation, one that is uniquely manifested by (some possible) persons and is inherently goal-directed and nondeterministic. It is not directed to any particular effects, but instead confers upon an agent a power to cause a certain type of event within the agent: the coming to be of a state of intention to carry out some act, thereby resolving a state of uncertainty about which action to undertake.’’[6]

But this relatively simple picture quickly gets muddier as O’Connor further develops his theory. It seems unlikely that we cause our actions randomly. We usually have reasons for our actions, such as desires, long term plans or moral principles. Many philosophers believe that these reasons should play a causal role during the creation of our actions.[7] O’Connor disagrees. The agent is the origin of action, not the agent’s reasons. The influence of reasons is explained through the way they structure how we can act. When we consider acting, we usually have a limited set of options. If I am hungry, then I might eat some bread or fruit. But I will not take a nap. Taking a nap does not come to mind as a serious option because I have no reason to take one. My reasons structure my action in such a way that I want to eat either bread or fruit, but leave me without any interest in a nap. In this way, reasons limit and structure my options for action. These reasons are often not equal in strength. Perhaps I dislike fruit and strongly prefer the taste of bread. In that case, I realise that I could eat some fruit and consider it as an option, but I have a strong urge to opt for bread instead. I have multiple options because I have multiple conflicting reasons. I now get to decide, as an agent, what action to take. In most cases, I will eat some bread because I have stronger reasons to eat bread than to eat fruit. But I still consider fruit an option and in some cases I decide to go for fruit. Which scenario actually comes about depends fully on the agent. The reasons do not cause anything directly, they just form a non-causal propensity to act in a certain manner.[8]

There are cases where some reason is so overwhelmingly strong that all other options disappear. In those cases, we cannot freely choose what to do since we only have one option. If somebody offered me the choice between the death of my entire family or a free chocolate bar, then I do not really have a choice. Murdering my family is against everything I am as a person, whereas accepting free chocolate agrees with my personality and beliefs. I am, in a sense, determined to pick the chocolate bar because choosing death does not come to mind as a viable option. But, O’Connor contends, these situations are rare. I often do have multiple options backed by different reasons and I often have to choose.[9]

Apart from their structuring role, reasons are also used by O’Connor to explain our actions. In some theories, the reason that causes the action is used as an explanation for the action. This fairly natural approach is unavailable to O’Connor, since reasons cannot cause free actions in his view. Instead, he gives the following account:

‘’If an agent acted in order to satisfy his antecedent desire that X, then:

---

[5] O’Connor, *Persons and Causes,* 108-125.

[6] Timothy O’Connor, ‘’Agent-Causal Theories of Freedom,’’ in *The Oxford Handbook of Free Will (Second Edition)*, ed. Robert Kane (Oxford: Oxford University Press, 2011), 313.

[7] O’Connor, *Persons and Causes,* 86-91.

[8] O’Connor, *Persons and Causes,* 91-101.

[9] O’Connor, ‘’Agent-Causal Theories of Freedom,’’ 319-321.

1. prior to this action, the agent had a desire that X and believed that by so acting he would satisfy (or contribute to satisfying) that desire;
2. the agent's action was initiated (in part) by his own self-determining causal activity, the event component of which is the-coming-to-be-of-an-action-triggering-intention-to-so-act-here-and-now-to-satisfy-X;
3. concurrent with this action, he continued to desire that X and intended of this action that it satisfy (or contribute to satisfying) that desire; and
4. the concurrent intention was a direct causal consequence (intuitively, a continuation) of the action-triggering intention brought about by the agent, and it causally sustained the completion of the action.''[10]

For some desire *X* to be the reason-explanation of the action, the agent should have that desire the entire time the action is performed. The agent should believe that the action will help satisfy *X* and the agent should be the origin of the action through her agent-causal powers. But it is most important to note that *X* should be referred to in the content of the action-triggering intention. Any desire *Y* that the agent has and that can be satisfied through the proposed action, but is not referred to in the content of the action-triggering intention, is not a reason that the agent acted for. All desires like *Y* cannot, in O'Connor's view, serve as an explanation for the action.

In a nutshell, O'Connor's theory is that the agent has agent-causal powers that can cause an action-triggering intention. This intention in turn causes the action. Because of these powers, the agent is in control and is free. But the decisions of the agent are structured by her reasons. Some options are more alluring than others and some options are not considered at all. Our freedom is limited. But reasons do allow us to explain the agent's behaviour. The desire or reason that is referred to in the content of the action-triggering intention can serve as an explanation of the action.
In the coming chapters, we will see whether this theory is coherent, whether it can successfully explain our freedom and our relationship with reasons and how well the theory fits with empirical data.

[10] O'Connor, *Persons and Causes*, 86.

II Coherence

Some philosophers are of the opinion that the above story is incoherent. Specifically, the concept of agent-causation is supposed to be incoherent. I will discuss two attempts to show this. The first attempt is formulated by Galen Strawson.[11] The second attempt concerns datability. I will deal with these attempts in turn, seeing whether they are justified or not.

According to Strawson, how one acts at moment *T1* depends on how one is at moment *T1*, mentally speaking. We might refer to this ''how one is at moment *T1*'' as one's mental state at that time, or *M1*. Let us call the action *A*. Thus, *M1* at *T1* determines[12] *A*. *M1* may include whatever we want it to include. We might include the agent's emotions, beliefs, her deliberations at or just before *T1*, *et cetera*. But if we want to claim that action *A* is a free action, then we have a problem. For *A* to be a free action, it should have been the agent herself who decided that *M1* would be the case. If the agent did not choose *M1*, and *M1* determined *A*, then it was not because of the agent that *A* occurred. Thus, if we are to defend the possibility of freedom, we have to conclude that the agent chose *M1*. But how did the agent make this decision? Let us say that she makes decision *D* at time *T0*. *D* is the decision that her mental state will be *M1*. This decision cannot come from nowhere. The agent's mental state at *T0*, *M0*, provides the background from which *D* springs. As said before, how one acts, or decides, depends on the agent's mental state at that time. The same problem as before appears. If the agent chose *D* freely, she must have chosen *M0* as well. The choice of *M0* at time *T-1* requires *M-1* however, thus necessitating the choice for *M-1*, which depends on *M-2*, *et cetera*. We discover an infinite regress of choices. Every choice or action requires a mental state, which requires a choice, *ad inifinitum*. The only way out is to allow the agent to create her mental state *M1* freely, without the support of some prior mental state *M0*. This would come down to allowing the agent to be a *causa sui*, something that can create itself out of nothing. But concluding that every agent is a *causa sui* is completely unacceptable, to most philosophers anyway. The idea that an agent could create itself out of nothing, completely independent on her character, history or cultural and social circumstances, just does not seem in any way plausible. Strawson concludes that that agent causalists[13] are now caught between an infinite regress and the acceptance of the agent as *causa sui*. Neither option results in a plausible theory, thus terminating the viability of agent-causal theories of free will.[14]

O'Conner partly agrees with Strawson's insistence on the importance of mental states: our reasons structure our choices. Because of how we are and the situation at hand, we have a number of options. Some options are more likely to be chosen than others, again because of our mental state and the situation. In this structuring role lies the power of the mental state. Up to this point, Strawson is right in pointing out how mental states are important for determining action.[15]
This apparent harmony between O'Connor and Strawson is short-lived. According to O'Connor, the final choice is made by the agent herself. This happens independently of *M* or anything else.

---

[11] Galen Strawson, *Freedom and Belief*, (Oxford: Clarendon Press, 1986), 25-31.

[12] What kind of determining this is supposed to be is unclear. Strawson also uses terms such as *A* because of *M1* and *A* as a function of *M1.* I have chosen to consistently use the term 'determining', though I will not necessarily mean causal determination. See Strawson, *Freedom and Belief*, 27-29. And Galen Strawson, ''The Bounds of Freedom,'' in *The Oxford Handbook of Free Will (Second Edition)*, ed. Robert Kane (Oxford: Oxford University Press, 2011), 442-444.

[13] Strawson believes that his argument can be used against all available libertarian theories of free will. I focus here only on the impact of his argument on O'Connor's theory.

[14] Strawson, ''The Bounds of Freedom,'' 442-449.

[15] O'Connor, ''Agent-Causal Theories of Freedom,'' 319-322.

Strawson would object. How can the agent make this final choice? There is nothing about the agent that could create, cause or explain the final decision. According to Strawson's starting premise, there must be some *M* to determine the final choice. Otherwise, we must accept the agent as a *causa sui*, creating its own *M*.[16]

O'Connor would answer that no *M* is required for the final decision. *M* has already been given due importance through its structuring role. The final decision is to be made by the free agent. It might be objected that if it is not *M* that determines the final choice, then what is it? O'Connor can simply reply that it was the agent that determined it, using her agent-causal powers. There is no more to this story. No *causa sui* powers are necessary, since the agent does not need to create itself (in the form of a *M*) to use those powers.[17] To say that there absolutely must be some *M* at work here, either *M* created by the agent as a *causa sui* or *M* as existing naturally, requires an extra argument. Strawson does not provide this argument, saying: ''I take it [the premise that *M* determines action] to be incontrovertible, quibbles aside, and will not defend it''[18].

I think that O'Connor's defence is sufficient to refute Strawson's claims. But it does leave it mysterious on what grounds the agent makes her final choice. We are told that there is nothing about the agent or her reasons that determines the action. So, if reasons are out of the picture at the final decision, then on what grounds does the agent decide? It seems that the choice is made out of the blue as it were. This, apparently, is just something the agent can do, through her powers. This is not an incoherent idea, but as an explanation of free action it is not highly satisfactory. We would like to know how the agent can decide between two options without fundamentally relying on her reasons. The only explanation we get is that she just can. That, as explanations go, is not highly informative.

The second attempt to show the incoherence of agent-causation involves, as said, datability. Every event, including a human acting in some way, happens at some specific moment. For example, some action *A* might occur at time *T1*. Philosophers might wonder why *A* happened at *T1*. Why not at *T2* or *T3*? What determines when *A* happens? Some have claimed that an agent-causalist cannot answer these questions. This is because actions are caused by the agent, an enduring object and not by an event, which is a dated occurrence.[19] Ginet formulates it in the following way: ''More decisive is the difficulty, pointed out by C. D. Broad, that if the cause of the mental occurrence is just me, just the enduring entity, and no event at all, then it cannot explain what it needs to explain. A merely enduring thing as cause lacks the features needed to make it capable of explaining the particulars of the mental occurrence. It cannot, for instance, explain its timing. The mere fact that I was there cannot explain why this mental act occurred just when it did rather than earlier or later, when I was also there.''[20]

O'Connor responds by pointing out that this is not a problem at all. Some agent at time *T1* had the power to cause *A*. Because the agent is free, she could decide to either do it or not do it. *A* simply occurred at the exact moment the agent chose to let it occur. There does not seem to be a straightforward reason why the agent has to be a datable entity to allow this to be possible. O'Connor further supports this claim by referring to structuring reasons. Sometimes we have reasons to cause *A*, sometimes reasons to cause *B*, *et cetera*. Us having certain reasons at some specific time can further show why we caused some event at that specific time.[21]

---

[16] O'Connor, ''Agent-Causal Theories of Freedom,'' 319-322.
[17] O'Connor, ''Agent-Causal Theories of Freedom,'' 319-322.
[18] Strawson, ''The Bounds of Freedom,'' 445.
[19] Carl Ginet, *On Action,* (Cambridge: Cambridge University Press, 1990), 13-16.
[20] Carl Ginet, *On Action,* 13-14.
[21] O'Connor, ''Agent-Causal Theories of Freedom,'' 614-617.

I agree with O'Connor that there is no particular reason why it would be incoherent for an agent to cause some event at a specific time. But I do not think that the above explanation of the timing of action is satisfactory. O'Connor merely says that the agent has the power to cause events to occur at specific times. We are not told how the agent can do this or why the agent chooses to act when she does. Pointing towards structuring reasons to explain timing seems insufficient, since it is the agent herself that makes the final decision. It would be strange to claim that some action is caused by some agent, except for the timing of the action. Apparently, the agent can choose to act at whatever time she wishes without further reasons to do so. Nothing more is added. This solution is not very satisfying nor very informative. We still do not know how or why the agent determines the timing of action, she just does.

To say, with Ginet[22], that agent-causation is incoherent because it has trouble explaining timing would go too far. It is not the case that O'Connor cannot explain timing at all. It is just that his explanation is not very informative and that a richer explanation would be preferable. This is a downside to O'Connor's theory, but a lacking explanation does not warrant the label of incoherence. Still, it does not strengthen O'Connor's position to have uninformative explanations as a part of his theory. O'Connor's strategy to negate this negative conclusion is to claim that he does not have to explain the precise timing of the free action at all.

''Let us recast the objection in this way: in the agent causation view, a feature of, or fact about, *e* – that it occurred at *t* rather than at *T1* – is a result of some causal factor that fails to explain it. And isn't that puzzling? If that is the problem, it is also a problem facing the idea of indeterministic event causation.''[23]

O'Connor then notes that an indeterministic event can have multiple outcomes and that there is no further reason why one outcome occurs rather than some other. If indeterministic events get away with this, then why not agent causal events?[24]

It is indeed true that the outcome of some indeterministic events cannot be analysed beyond pointing to chance. This lack of a further, more satisfying explanation is indeed a downside to the concept of chance. We introduced chance to explain certain events that seemed indeterminate, such as radioactive decay (RD). This introduction allows us to explain RD up to a certain point, but we reach this point rather quickly. If we ask why a particular piece of radioactive matter decayed in such and such a way, we can only answer that this is because of chance. But further questions cannot be answered. For example, how did chance select between several possible manners in which the radioactive matter could have decayed? The only thing we can say here is that chance simply does select, somehow. There is not much more we can say about it. But that explanation, ''chance just behaves like that'', is rather uninformative. Thus, I would agree with O'Connor that the explanations the concept of indeterministic event causation provides can be uninformative.

Then again, the vice of the uninformative explanations is compensated by other merits the concept of indeterministic events has. For example, explaining radioactive decay without chance would be difficult if not impossible.

O'Connor could follow a similar tactic and claim that the concept of agent-causation is necessary to explain free human behaviour. In fact, he does argue that alternative theories that do not use the concept of agent-causation are worse at explaining free human behaviour.[25] I will not check whether this claim is true, since I would have to deal with all currently available theories of free will. I simply do not have the space to do so. But I will note that even if we need agent-causation to make sense of

---

[22] Ginet, *On Action,* 14.

[23] O'Connor, *Persons and Causes,* 76.

[24] O'Connor, *Persons and Causes,* 74-76.

[25] O'Connor, *Persons and Causes,* 23-42.

free action, its uninformative nature is still a downside. We would accept this negative quality, since we do not really have a choice. But it would be a downside of agent-causation nonetheless.

We have seen here that O’Connor’s theory is not incoherent. Strawson and Ginet underestimated the versatility of agent-causal powers. However, the discussions here did show some signs that O’Connor has trouble to properly explain how certain aspects of agent-causation and human action work. It remains unclear how the agent can make the final decision without relying on something about herself. What does she base her decision on? The timing of action also remains poorly explained. Thus, this is a chapter of mixed blessings for O’Connor. I found no reason to mark him as incoherent, but I did note that some of the explanations he gives are unsatisfactory.

III Explaining Freedom of Action

The main goal of O'Connor's theory is to make it clear how agents can be free. O'Connor's proposal is to explain this phenomenon through agent-causation. This primitive power is supposed to grant the agent control over her action. But some have claimed that O'Connor's attempt fails. An important motivation for this claim is the luck objection. The luck objection claims that agent-causal freedom comes down to mere luck. If this is true, then O'Connor has failed because mere luck is clearly not sufficient for freedom. I consider here two forms of the luck objection. One of these is Van Inwagen's famous rollback argument[26]. But I will start with the luck objection Alfred Mele formulated.[27]

Imagine Caesar standing at the bank of the Rubicon. He can either cross it and march to Rome or retreat and remain in Gaul. The former option leads to civil war, the latter to his personal demise. Caesar's personality and motivations are such that the reasons for crossing or retreating are equally strong. In our world[28], Caesar decided to cross the Rubicon, plunging the Roman Republic into a bloody civil war. But, if O'Connor is correct, there is some possible world in which Caesar chose not to cross the Rubicon. This world is, up to the point where Caesar decides to cross, completely identical to ours. Caesar has exactly the same reasons, motivations, personality *et cetera*. And yet, his decision is the opposite of the decision of our Caesar. Mele now asks how this can be anything but chance. Two identical people are in the exact same situation, but respond differently. There appears to be no reason why this is so. Dumb luck seems to be the only way to account for the difference.[29]

O'Connor's answer to Mele's objection is fairly straightforward. Caesar's decision is indeed not caused by his personality or motivations. Therefore, there could be a possible world with the exact same Caesar as we had in ours, but who decided not to cross the Rubicon. But this difference is not a consequence of mere luck or chance. It was Caesar's freedom that caused this split. He, as an agent, had the power to make a decision independently. This independent freedom is compatible with there being different worlds in which Caesar makes a different decision. That is what freedom is after all, the possibility to do something else[30].[31]

I consider this reply sufficient to refute Mele's version of the luck objection. O'Connor shows that Caesar's behaviour does not necessarily rest on mere chance, but is perfectly understandable in an agent-causal framework.

Peter van Inwagen has developed a, what he calls, ''plausible, intuitive version''[32] of the luck objection. Again, Caesar stands at the Rubicon, contemplating whether to cross or not. *T1* is the moment in which he makes his fateful decision to cross. As he and his army are crossing, God interferes. He rewinds the universe to moment *T1*. Caesar needs to make his choice anew, in the exact same situation as before. But, as we have seen, his decision is not determined by the situation, but only by Caesar's agency. Thus, Caesar might make a different choice this time. Let us now say that God repeats this rewinding over and over again. At a thousand rewinds, He stops and looks at the results. Since Caesar's two options are equally strong, God will notice that in roughly half the

---

[26] Peter Van Inwagen, ''Free Will Remains a Mystery,'' *Philosophical Perspectives* 14 (2000), 1-14.

[27] Alfred Mele, ''Ultimate Responsibility and Dumb Luck,'' *Social Philosophy and Policy* 16 (1999), 274-293.

[28] Note that I am not claiming that my rendering of Caesar crossing the Rubicon is correct. It is merely an example. For example, it is not very likely that Caesar's choice at the Rubicon, if such an event even occurred at all, was a fifty-fifty choice.

[29] Mele, ''Ultimate Responsibility and Dumb Luck,'' 275-277.

[30] Compatibilists would disagree of course. But to O'Connor, multiple choices are necessary for freedom.

[31] O'Connor, ''Agent-Causal Theories of Freedom,'' 320-323.

[32] Inwagen, ''Free Will Remains a Mystery,'' 16.

cases Caesars crosses and in the other half he does not. Upon further experimentation, this 50/50 ratio turns out to be an accurate way to describe Caesar's behaviour. If we see such an obvious pattern, are we then not justified in believing it a matter of chance rather than free choice?[33]
It might seem that O'Connor can reply in the same way as before.[34] Caesar's behaviour was not determined by chance. Through his agent-causal powers, he determined his own actions freely. The recognisable pattern is merely a consequence of this freedom and Caesar's reasons, desires, *et cetera*. Just because there is a pattern does not mean that Caesar cannot be free.
But Van Inwagen anticipated this response and reacted to it in the very same essay. Even if agent-causation is what determines the decision, then we still cannot escape the impression that it is a matter of chance. Let's say that Caesar gets to choose whether to cross the Rubicon or not another thousand times. The 50/50 ratio persists. How is the knowledge that agent-causation is involved going to undermine the strong impression that Caesar's behaviour is governed by chance?[35] Van Inwagen puts it in this way:

''Nothing we could possibly learn, nothing God knows, it would seem, should lead us to distrust our initial inclination to say that the outcome of the next replay will be a matter of chance. If this much is granted, the argument proceeds as before, in serene indifference to the fact that we are now supposing Alice [Caesar in our example] to be the agent-cause of various sets of cerebral events that are antecedents of the bodily movements that constitute her acts.''[36]

I think the proper answer here is twofold. The first element is the familiar O'Connorian response. We need to realise that it is the agent that controls which decision gets made. That is her primitive power, to determine her own action. Some inclination to interpret the agent in a chancy way is not sufficient to seriously doubt her freedom.
This first element is clearly not enough to convince O'Connor's opponents, since they continue to have their chancy inclination. We can take away this inclination through explaining this feeling in a way that is compatible with the free agent as imagined by O'Connor. Take Caesar again and presume that he has agent-causal powers and is truly free. He is structured by his reasons in such a way that he has two equally strong options. Through his agent-causal powers, he makes a choice. If he had to repeat his choice over and over again, he would sometimes cross and sometimes not. This is simply a natural consequence of his freedom. Now invent a collection of observers with the power of rewinding the universe like God did some paragraphs before. When they observe, they merely see Caesar's behaviour. They soon notice that Caesar's behaviour at the Rubicon follows a clear pattern. Van Inwagen contends that the observers should conclude that Caesar is governed by chance. After all, this is all they can tell by Caesar's behaviour. This conclusion, I would say, is justified. To them, it does appear to be mere change. But they fail to realise that this pattern is a consequence of Caesar's freedom. For every decision, it holds that Caesar, in his mind, makes a conscious choice. Chance does not determine the behaviour, agent-causal powers do this. The chancy appearance is a consequence of freedom. The observers do not see this because they only see the appearance and not the power-based internal processes. Those powers only play their role within the mind of the agent. One cannot see someone else make a decision.[37] Sure, you could see another behave in a contemplative manner,

---

[33] Van Inwagen, ''Free Will Remains a Mystery,'' 12-17.
[34] He does in fact. The reply he gives to Mele was also meant for Van Inwagen.
[35] Van Inwagen, ''Free Will Remains a Mystery,'' 16-17.
[36] Van Inwagen, ''Free Will Remains a Mystery,'' 17.
[37] Perhaps one could with the right knowledge and a sophisticated brain scan. But answering that question properly would go too far for the purposes and scope of this thesis. Observing will here be limited to the use of a regular human eye.

indicating that he is thinking. Then you might observe him getting up and start doing something. A decision has been made, you might conclude. Still, you only saw behaviour, not the decision making itself. The decision and the accompanying powers have their place in the mind. That is where they are active.[38] The invisibility of the agent-causal powers has led the observers to base their conclusions on the chancy appearance of the actions.
What should we take away from this lengthy story? First, that free behaviour can seem chancy to outsiders. Second, that this appearance is a consequence of freedom and structuring reasons. The first conclusion allows us to agree with Van Inwagen that, in some specific circumstances, there is some more or less natural inclination to see free behaviour of the agent-causal variety as chancy. The tale of Caesar is an example where such an inclination may arise. The second conclusion allows us to disagree with Van Inwagen that this inclination spells trouble for the agent-causalist. If we learned that agents had agent-causal powers that were hidden from normal observation, we would conclude that the chancy appearance was misleading. Underneath this appearance, there was true freedom.
Perhaps an example given by Van Inwagen can bring the point home. A friend of mine is running for public office and I know some secret about him. This secret would ruin his campaign and with good reason. This friend begs me to not tell anybody. I am torn between doing my duty as a citizen and my close ties to my friend. The chance of me telling on him is 46 percent whereas the chance of me remaining silent is 54 percent. Van Inwagen now contends that I cannot possibly promise him in good faith to remain silent. There is a 46/100 chance that I will tell on him. How could I ever promise my friend silence when there is a good chance that I will speak up? This result shows that I am not truly free. If I were truly free and in control of my actions, then I could make promises to my friend.[39]
I think that the agent-causalist can explain, in his own terms, what is going on here. The agent, in this case me, has to power to make the final decision. I can remain silent or I can spill my friend's secret. The only reason why I would not be able to promise anything is because I am uncertain what I want to do. I have strong reasons to tell and strong reasons to remain silent, so the choice is tough. Because of this uncertainty, I cannot make promises about what I will do. Thus, Van Inwagen is right about my inability to promise. But he is mistaken in his conclusion that my inability to promise undermines my freedom. The problem is not that I cannot act freely, the problem is that I simply do not know yet how I *want* to act. If I made up my mind and determined what I wanted, then I could make a decision and make promises. In the scenario sketched by Van Inwagen, this simply has not happened yet. I am still deliberating, which is why I might act either way. At the end of my deliberation, I will make my final and free decision to either remain silent or speak up. If I decide to remain silent, I can inform my friend of my decision and safely make promises to him. The promise is in good faith, because I know that I am in control of my action and because I want to remain silent. And with that, I have shown that the example given by Van Inwagen does not, in any obvious way, undermine O'Connor's theory. Without modifying the example or O'Connor's theory, we could make Van Inwagen's scenario fit with O'Connor's views.

We have seen that Mele's and Van Inwagen's objections fail. Their objections depended on the idea that free action as described by O'Connor can, in certain circumstances, look like behaviour that is controlled by chance. I agreed that this is true, but appearance is not enough to refute O'Connor's claims. We are interested in what happens inside the head of the agent, not in what her behaviour looks like to outsiders. I tried to prove this by describing the given examples in a way that fits with

[38] Agent-causal powers are primitive, so what we can say about them is limited. Still, I think that what I say here would not be contradicted by O'Connor.
[39] Van Inwagen, ''Free Will Remains a Mystery,'' 17-18.

O’Connor’s theory. In this I was successful. The claim that O’Connorian freedom is equal to or indistinguishable from mere luck is thereby refuted.

IV Explaining the Relation between Reasons and Action

When agents act, they act for a reason. This reason can be a desire, a moral principle, a practical consideration and so on. To integrate this fact in O'Connor's theory is not trivial. This is because O'Connor thinks that reasons cannot cause actions, or action-triggering intentions, directly. Thus, O'Connor needs to find a different, indirect route to explain the relation between reasons and action. He does this by claiming that reasons can influence the agent, without determining the action. This influence is, according to O'Connor, significant enough to provide reason-explanations for action. If Billy eats a sandwich, then we can justifiably count, say, his hunger as an explanation. This also holds if hunger is not the direct cause of Billy's action.[40]
Still, some philosophers have claimed that O'Connor's account of the relation between reasons and action is inadequate. I will consider three sources of complaints: by Randolph Clarke[41], by Richard Feldman and Andrei A. Buckareff,[42] and finally by Niels van Miltenburg.[43] The aim of this chapter is to find out whether these complaints have any merit and how destructive they are to O'Connor's theory.

Let us first quickly repeat O'Connor's theory of acting for a reason:

''If an agent acted in order to satisfy his antecedent desire that X, then:
1. prior to this action, the agent had a desire that X and believed that by so acting he would satisfy (or contribute to satisfying) that desire;
2. the agent's action was initiated (in part) by his own self-determining causal activity, the event component of which is the-coming-to-be-of-an-action-triggering-intention-to-so-act-here-and-now-to-satisfy-X;
3. concurrent with this action, he continued to desire that X and intended of this action that it satisfy (or contribute to satisfying) that desire; and
4. the concurrent intention was a direct causal consequence (intuitively, a continuation) of the action-triggering intention brought about by the agent, and it causally sustained the completion of the action.''[44]

Or more simply put, the agent desired *X* and caused an action-triggering intention, which in turn event-caused the rest of the action. During the action, the agent continued to desire *X* and the agent must consistently be the origin of the action through his action-triggering intention. Note that O'Connor claims that we can only say about a desire or a reason that we acted because of it or that it explains our action if the desire or reason is part of the content of our intention. So only if our intention is 'to do action *A* in order to satisfy reason *X*' can we say of *A* that *X* is the reason we did it.

Clarke objects to the above account. To illustrate his objection, we could imagine Otto von Bismarck in July 1870.[45] He had two desires that interest us. Bismarck wanted his country, Prussia, to go to war against France and he wanted to unify Germany.[46] Let us say that his desire for war caused Bismarck

[40] O'Connor, *Persons and Causes,* 85-86.
[41] Clarke, *Libertarian Accounts of Free Will,* 138-144.
[42] Richard Feldman and Andrei A. Buckareff, ''Reasons Explanations and Pure Agency,'' *Philosophical Studies* 112 (2003), 135-145.
[43] Van Miltenburg, *Freedom in Action,* 157-162.
[44] O'Connor, *Persons and Causes*, 86.
[45] As with Caesar, I do not claim that this tale about Bismarck is historically correct.
[46] Bismarck has these desires independently. He seeks a war against France even if this would not bring unification closer and vice versa. It is not the case that these desires are completely unrelated, but it is important to note that the neither desire is dependent on the other desire for its existence.

to insult France. This insult was supposed to, and did, cause the French to declare war on Prussia. But let us suppose that the content of Bismarck's intention was 'to insult France, in order to unify Germany' and not 'to insult France, in order to go war against them'. After all, Bismarck hoped that the insult would contribute to his efforts for unification. O'Connor would have to point to Bismarck's desire for unification as an explanation for his insult to France. That is the desire referred to in the content of the intention. But it was Bismarck's desire for war that caused him to insult France. Thus, it seems more sensible to say that it was his desire for war and not his desire for unification that should serve as the explanation for his action. We might say that the desire for unification is still relevant, but the desire for war, which is the cause of the insult, seems much more important. The fact that O'Connor must choose the former desire instead of the latter can therefore be seen as a failure of his theory.[47]

Clarke does express the worry that the above might not be a valid criticism of O'Connor. He says that Bismarck's desire for war causes an action-triggering intention, but that it does not activate any agent causations. Only event-causations take place. This fact might undermine the example. One could also worry whether Bismarck's desire for war could cause an intention that refers to his desire for unification. Perhaps desires cannot influence us to adopt an intention that refers to a different desire. Clarke claims, without further proof, that this second worry is unnecessary. Desires can influence us to adopt intentions that refer to a different desire in cases of self-deception.[48]

I think that Clarke's first worry is valid and fully undermines the significance of the example. It is possible for a desire or reason or something else about our mental state to cause an action without the use of agent-causal powers, according to O'Connor. This happens, for instance, when we act out of habit, e.g., when we are breathing. We usually do not make a conscious decision to breath. We breath automatically. But such automatic actions are not free.[49] Only actions with an agent-causal origin are truly free in O'Connor's view.[50] And in this case, we are investigating the relation between free action and reasons. Thus, the objection given by Clarke does not work against O'Connor because it describes an unfree action and not a free one. O'Connor can simply reply that his reason-explanations account only holds for free action.

Nevertheless, I think that the example given by Clarke can easily be modified to be acceptable as a free action, as imagined by O'Connor. I will give such a modified story. At the same time, I show that the second worry can indeed be solved by considering cases of self-deception. Clarke did not work that idea out properly, so I will do so here.

Imagine a man, call him Abraham Aaronson, walking through London late at night. Suddenly, he hears a shout and sees another man lying in the Thames. This man is Bernard Bluestockings, his arch-nemesis. Bluestockings is drowning and nobody, other than Aaronson, is nearby to save him. Aaronson can now either let him drown or he can save him. Aaronson has, he thinks, a moral dilemma before him. Bluestockings is an evil man and the world would be a much better place without him. On the other hand, allowing a fellow human to drown when he could be saved is very wicked, close to murder. The two options for acting are equally strong. The choice is in that sense similar to the case of Caesar at the Rubicon.

But Aaronson is mistaken in his reasons for the dilemma. In reality, his dilemma is purely selfish. He wants Bluestockings to drown because he hates him and wants him out of his life. Then again, saving a drowning man would be honourable and Aaronson is strongly attached to the image of himself as a

[47] Clarke, *Libertarian Accounts of Free Will,* 140-141.

[48] Clarke, *Libertarian Accounts of Free Will,* 141-142.

[49] Of course, we could make the conscious decision to breath, thus making it a free action. But in most cases, we automatically start to breath.

[50] O'Connor, *Persons and Causes,* 68-74.

man of high principles. The dilemma is a battle between hatred and vanity, not between opposing moral principles.
Let us say that Aaronson ultimately chooses, through his agent-causal powers, to save his rival. The true motivation for this is that it will make Aaronson feel morally superior. But Aaronson thinks that his principles, and not potential gain, motivated him. Thus, it seems plausible that his intention was 'to save Bluestockings, because of my moral conviction that this is the good thing to do'. After all, that is how Aaronson thinks about this situation. It would make sense if this was reflected in the intention he decides to cause. But his moral conviction, which he truly had, was not at all the reason of his rescue attempt. Thus, it would be strange if his morality served as the explanation of his action. But if we look at O'Connor's conditions, then this would be the result. He had his moral conviction throughout the action and it featured in the content of the intention. We are also dealing with a case of free action. Aaronson can freely choose between the life or death of Bluestockings, just as in any regular choice. O'Connor's theory simply seems to fail here.
Perhaps O'Connor could counter that the vanity of Aaronson could not tend toward the intention 'to save Bluestockings, because of my moral conviction that this is the good thing to do'. Perhaps Aaronson's vanity could only motivate an intention with a content like 'to save Bluestockings, because of my vanity'. But this would be strange, considering that Aaronson is not aware of his own vanity. If his vanity were to feature in the content of his intention, would he then not be aware of it? O'Connor could go on to claim that Aaronson simply does not know the content of his own intentions. He has claimed that this is possible.[51] If the above is true, then the example I have provided is not actually possible. The failure of O'Connor's theory to properly deal with the Aaronson case would then no longer be a problem.
But such a solution is not very attractive. The creation of intentions is the source of our control over our actions. If we follow O'Connor and create this rift between intentions and ourselves, then our means of controlling our actions would become suspect. To control our actions through intentions whose content we are not aware of and apparently do not have full control over is rather suspicious. I do not think that such an account would be incoherent *per se*, but it certainly is not an elegant solution. I highly doubt that anybody would opt for such a theory unless there is absolutely no other alternative. Thus, saying that we are not aware of the content of our own intentions does not seem a promising route to deal with the Aaronson case.
Another way to solve the Aaronson problem would be to simply bite the bullet. O'Connor could maintain that Aaronson's moral fibre, and not his selfish reasons, are to explain his action. After all, Aaronson made his decision based on his moral convictions. Why not just accept these convictions as an explanation?
The example I have given can easily be adapted to make this reply seem implausible. Let us say that Aaronson's selfish reasons were incredibly strong, almost overpowering. His moral convictions barely managed to influence his decision at all. Their influence, though not zero, was negligible. We could even imagine a parallel and identical case, with the only difference being that Aaronson's selfish vanity was absent. His lust for revenge against Bluestockings would blow away his minor moral scruples and he would not save his enemy. To maintain, despite this difference in strength, that it was morality that explains Aaronson's action and not his vanity, seems completely implausible. I do not see how morality, with its near-zero influence, could ever explain the action. Biting the bullet does not seem to be a viable option. Another way out would be preferable.
One last alternative route would be to argue that Aaronson's action was not free. He was unaware of his true reasons. He was thus unable to make a fully informed decision. He chose blindly, not knowing why he made his decision.

[51] O'Connor, *Persons and Causes,* 87-88.

Intuitively, there is some merit to this claim. Aaronson was driven on by forces he did not recognise. This made it harder for him to criticise the validity of these motivations. Perhaps, if he realised that it was his vanity that pushed him to save Bluestockings, he would have acted differently. Or in the opposite scenario, if he found out that only his hatred kept him from saving his enemy, he might have disregarded that reason with disgust, thus favouring a rescue attempt. But these considerations were impossible because of his lack of self-awareness. This left less room for a properly thought out choice.

Still, despite his limited knowledge of his own motivations, Aaronson was still a free man. He was in control of the decision, since neither vanity nor hatred was strong enough to overwhelm Aaronson. His agent-causal powers, and not his reasons, determined his action. He also clearly had multiple options. His lack of information limited him somewhat, but not enough to say he was not free. As O'Connor said himself, freedom comes in degrees.[52] Aaronson might have been less free in the drowning scenario than he normally is, but I do not see any strong argument why he was decidedly not free.

All in all, I cannot see a straightforward way for O'Connor to dismiss the Aaronson case. The case also dodges Clarke's worries. If my story is plausible, then self-deception can indeed allow a reason to influence an agent to create an intention which, in its content, refers to some other, unrelated reason. And because this first reason influences the agent, rather than causing the intention directly as it had in Clarke's example, it does not run into the worry that my example is incompatible with an agent-causal free action. I conclude that O'Connor's account of reason explanations does not work properly in cases of self-deception, such as my example.

Feldman and Buckareff, as presented here, are mainly concerned about the content of our action-triggering intentions. Imagine an agent with two desires, *D1* an *D2*. *D1* is 'I want a beer', *D2* is 'I want something to drink'. The agent decides, on the basis of *D1* and *D2*, to go to a nearby pub. She causes intention *1* to occur. Intention *1* is 'I will go to the bar in order to get a drink'. The intention only refers to *D2* and not to *D1*. So, O'Connor would be forced to accept that only *D2* explains the agent's behaviour. This despite the fact that *D1* was also a significant influence on the agent. It gets even worse when we imagine the agent causing the intention 'I will go to the bar'. Intention *2*, as we might call it, does not refer to either *D1* or *D2*. It now seems that neither *D1* or *D2* can explain the agent's behaviour. The consequence of both intention *1* and *2* are destructive to O'Connor's theory. It seems undeniable that *D1* and *D2* are both important to the agent's action. Denying one or both of these desires the status of explanation for the action can only be described as an error.[53]

One might retort that intention *1* and *2* are not possible in these cases. Perhaps, the content of our intentions always refers to the desires that were important in bringing those intentions about. But I cannot see any reason why this would be so, apart from the fact that this would be very convenient for O'Connor. Supposedly, we cause these intentions ourselves through our agent-causal powers. It would be strange if some automatic process filled in the content of our intentions. O'Connor might of course attempt to describe and justify such a view of our psyche, or find some other way to ensure that the content of our intentions always mentions the relevant desires or reasons. But, to my knowledge, O'Connor has not done this. The closest O'Connor has come to a justification of the described view is the following: ''agent causation is conceptually tied to the agent's having reason for acting. We might say that agent causation is a triadic relation.''[54] But claiming that reasons feature in the content of intentions in the way described here just because of close relations is not sufficient. There are many ways in which this close relation could be expressed. We require additional

[52] O'Connor, *Persons and Causes,* 101-107.
[53] Feldman and Buckareff, ''Reasons Explanations and Pure Agency,'' 137-142.
[54] O'Connor, *Persons and Causes,* 88.

arguments to show that reasons and intentions do behave as O'Connor thinks. O'Connor has not delivered those arguments.
Cases such as described by Feldman and Buckareff cannot be dismissed in any obvious way. We have also seen that, for these cases, O'Connor's theory is unable to pick the right reasons as explanations. This means that we now have discovered a second type of case in which O'Connor's theory fails.

The adapted example of Clarke and the examples of Feldman and Buckareff spell major trouble for O'Connor. His theory was supposed to show us which of the reasons the agent had can explain her actions. O'Connor attempted this by relying on the content of our intentions. This move can be seen in step 2, the most critical step of his account. But, as the above examples show, the account does not work properly. They showed that the content of our intentions does not always refer to a reason that is suitable as an explanation. As a consequence, O'Connor's theory fails to point out the correct reason-explanation. In other words, the theory completely fails to fulfil its intended purpose.

The last philosopher I will consider here is Niels van Miltenburg. His problem with O'Connor is fairly straightforward. How do reasons influence the agent? As we have seen, it is the agent, not her reasons, that cause action-triggering intentions to occur. But the agent is influenced by her reasons. O'Connor nowhere explains how this influence is supposed to work.[55] O'Connor does write that this influence is causal and probabilistic[56], but that does not seem sufficient. In what sense is the influence causal if the agent herself is uncaused and if the action-triggering intention is caused by the agent and explicitly not by her reasons? The exact meaning of the probabilistic element of this influence is also hard to make out, though less so. We have seen that for every decision there is some objective chance that the agent will perform either one action or the other. This objective chance exists because of the influence of our reasons, but it remains unclear how this influence can bring these chances into existence. Somehow, they structure our actions in such a way as to create this propensities in us, without taking away our freedom. But we are never told how our reasons achieve this feat.
O'Connor's claim that the influence of our reasons is causal and probabilistic is not nearly enough for an explanation. Too many central questions remain unanswered. The lack of an explanation for this absolutely central tenet of his theory is hard to swallow.[57]

In the last two chapters, we saw that O'Connor's theory could resist the problems that were brought up by his opponents. In this chapter however, things are different. Both the modified case of Clarke and the cases of Feldman and Buckareff show the inadequacy of relying on the content of intentions for the purpose of reasons-explanations. O'Connor will need to either give up that particular idea or adapt it. Alternatively, he could give some account of why the content of our intentions always refers to the relevant reasons for acting. This is not an easy thing to do, especially if we consider the fact that O'Connor needs to maintain that the agent is in control of her action. If there is some psychological or biological function that automatically fills in the content of our intentions, sometimes in such a way that we are not even aware of what the content of our own intentions is, then we will start worrying whether we are truly in control. I might have caused the intention to occur, but if I am not in control of its content, then how can I control the action that follows from my intention? I do not think that objections such as the above are strong enough to completely rule out the possibility of an account that states that the content of our intentions always refers to the

[55] Van Miltenburg, *Freedom in Action,* 160-162.
[56] O'Connor, *Persons and Causes,* 95.
[57] See Derk Pereboom for a similar criticism of O'Connor. Derk Pereboom, *Free Will, Agency and Meaning in Life,* (Oxford: Oxford University Press, 2014), 60-62.

relevant reasons. But the above objections do make it unlikely that a solution can be found that is in any way elegant.
Secondly, Van Miltenburg is right that O'Connor should give us some account of how reasons can influence us. And again, this account should not threaten our causal control over our actions. So, until O'Connor has given an account of how the content of our intentions is formed and how reasons can influence us, his theory on the relation between reasons and free action is inadequate.

V Empirical Data and Agent-Causation

O'Connor does not claim that his theory rests on a great amount of empirical data. In fact, he claims that the empirical sciences cannot, at the moment, confirm or deny the existence of agent-causation. They simply have not developed far enough yet to give the final judgement on the nature of our mind and our actions. Still, it would have been awkward if the empirical data we do have contradicted O'Connor's theories. That would leave O'Connor with the unpleasant task of having to explain this opposing data away. But O'Connor sees no reason why current empirical findings would contradict his theory. He admits that current findings do not support his theory in any interesting way, but they should not form a problem either.[58]

Whether O'Connor is right is a matter of debate. Some philosophers have thought that the current status of the empirical sciences is in fact highly problematic to deal with for agent-causal libertarians.[59] To deal with this thorny issue in a complete and adequate manner would force me to deal with complicated side-issues, such as the correct interpretation of neuroscientific data. That would go beyond the scope of this thesis. So instead of investigating the relationship between the sciences and agent-causation *in toto*, I will restrict myself to one particular type of data; phenomenological data.

Phenomenology is interesting here because it provides the only kind of data of which O'Connor claims that it supports his theory. He says that when we act, we feel in control of our action. We are not merely pushed around by our reasons, we make our own decision. ''Likewise, in the deliberate formation of an intention, the coming to be of my intention doesn't seem to me merely to occur at the conclusion of my deliberation; I seem to experience myself directly bringing it about.''[60] O'Connor thinks that this might point us in the direction of an agent-causal theory. Of course, it is no knock-down argument. Our experience might be mistaken or wrongly interpreted, but it can still serve as an indicator that agent-causal theories are on the right track.[61]

O'Connor never fully works out these thoughts on the phenomenology of free action. His remarks are simply based on his own experience and intuitions. Still, phenomenology remains interesting because it seems to be one of the few pieces of relatively concrete evidence for agent-causation, presuming O'Connor is right. A deepened knowledge of our experience could tell us whether O'Connor is right in presuming that the phenomenology of action is in his favour. The rest of this chapter is dedicated to sketching a way to investigate our phenomenology in a more systematic way than O'Connor has done. This sketch is based on the work of Eddy Nahmias et al.[62]. It is not meant to be detailed and it will not completely work out all the pros and cons of this type of research. The only use of the following sketch is to showcase some of the potential of a more systematic approach of phenomenology.

Settling phenomenological debates is not a trivial task. Phenomenology essentially rests on subjective feelings and this makes it hard to give formal arguments in favour of one's position concerning our experience. If O'Connor has one intuition about our experience and another philosopher has a different intuition, then how do we determine who is right? This is why Eddy Nahmias et al. suggest moving away from such subjective debates between professional philosophers. Nahmias claims that we should investigate our phenomenology in a more systematic way. Researchers could interview regular people about their experience or ask them to fill in forms.

---

[58] O'Connor, *Persons and Causes,* 108-126.

[59] See Pereboom for a discussion: Derek Pereboom, *Living Without Free Will*, 69-88.

[60] O'Connor, *Persons and Causes*, 124.

[61] O'Connor, *Persons and Causes,* 123-125.

[62] Eddy Nahmias, Stephen Morris, Thomas Nadelhoffer and Jason Turner, ''The Phenomenology of Free Will,'' *Journal of Consciousness Studies* 11 (2004), 162-179.

In this way, we could gather a large amount of phenomenological data. This data can then be used to create theories on what our experiences are like, generally speaking.[63]
Nahmias admits that this type of research is not exactly easy, since it is not straightforward what questions we should ask or how we should interpret the answers.[64] But he presumes that it should be possible to formulate questions whose answers can tell us something about our experience. If such questions are not possible, then we might wonder how phenomenology can serve as evidence at all for people like O'Connor. After all, how can it be that our experience is such that it can support a certain theory, but be completely unable to answers questions in order to show this? If our phenomenology truly has an agent-causal tendency, then this should show in the type of answers we give to certain questions.[65]
The type of research Nahmias has in mind will focus on the experience of laymen rather than trained philosophers. A problem one could have with this approach concerns the lack of expertise in the layman. It can be quite hard to discern and correctly label our experiences. Thus, it might be argued that only trained experts can investigate their mind properly. They possess a rich vocabulary to name their feelings and they have a lot of experience with introspection. Laymen lack both qualities, making it harder if not impossible for them to precisely discern their experiences.[66]
But are trained philosophers truly more desirable than laymen? Nahmias argues that this is not so. Their theoretical training might entail more detailed or precise reports, but it will also deform their experience. Because they are so deeply involved in philosophical work on the mind, they will tend to interpret their feelings in such a way as to fit their earlier knowledge and convictions. This gets worse if the introspecting subject defends some philosophical position herself. Take O'Connor for example. He believes in agent-causation and thus is more likely to interpret his experience to match his theories. He does not do this out of malice or to spread misinformation. The theory-ladenness of his observation influences his experiences. This is a normal and natural process, but it does mean that trained philosophers are less reliable to base one's research on.[67]
Of course, we might counter that laymen are also affected by their beliefs and their terminology. Laymen typically have certain ideas about free will and the workings of our mind. This will influence their reports. Therefore, the data we gather from them is not completely innocent either.
Nahmias seems to accept that there is some force behind this objection. Data provided by laymen will not be pure either. But the opinions and beliefs of laymen are far less precise and outspoken than those of professional philosophers. This gives us reason to think that laymen data is less tainted by philosophical convictions than those of professional philosophers, according to Nahmias.[68]

I did not attempt here to completely work out the relation between empirical data and O'Connor's work. What I did do was to suggest that were phenomenology is concerned, we should consider turning to systematic empirical research. This suggestion was based on a paper by Nahmias et al., whose proposal for further research I have attempted to sketch here. A number of reason were presented why a systematic approach to phenomenology might be better than the reflections of individual thinkers on their own experience. I do not pretend that I have been detailed enough to prove beyond a doubt the usefulness of systematic phenomenological research. I did not add anything significant to the ideas of Nahmias, nor did I criticize him in any deep way. Still, I hope I

[63] Nahmias, Morris, Nadelhoffer and Turner, ''The Phenomenology of Free Will,'' 164-165.
[64] See Nahmias and Turner for an example of this difficulty. Eddy Nahmias and Jason Turner, ''Are the Folk Agent-Causationists?'' *Mind & Language* 21 (2006), 597-609.
[65] Nahmias, Morris, Nadelhoffer and Turner, ''The Phenomenology of Free Will,'' 177-178.
[66] Nahmias, Morris, Nadelhoffer and Turner, ''The Phenomenology of Free Will,'' 170-174.
[67] Nahmias, Morris, Nadelhoffer and Turner, ''The Phenomenology of Free Will,'' 170-174.
[68] Nahmias, Morris, Nadelhoffer and Turner, ''The Phenomenology of Free Will,'' 171-172.

showed the potential of such research. We can use this potential to cast more light on our experiences while acting. That is important to this thesis, since O'Connor believes that phenomenology might serve as an indicator that his theory is right. By systematically investigating our experiences we can, hopefully, find out whether O'Connor's optimism is justified.

Conclusion

My goal here was to critically assess O'Connor's theory on human action and freedom. First, I looked at claims by Strawson and Ginet that agent-causal theories are incoherent. Strawson tried to show this by pointing to our mental state at the time of the decision and Ginet through describing the difficulties an agent-causal account has with datability. It soon turned out that things were not so bleak as Strawson and Ginet presented them to be. Both philosophers underestimated the possibilities of the concept of agent-causal powers.

Next, we saw that the objections of Van Inwagen and Mele, concerning the alleged freedom of the agent in O'Connor's theory, failed for similar reasons. It turned out that agent-causal powers could present the agent with freedom and did not deliver her to chance. This despite the intuitive appeal some of Van Inwagen's examples had.

But agent-causal powers did not help O'Connor overcome the problems we noted in chapter five. His account of reasons-explanations was not considered adequate. There were clear examples in which O'Connor could not point out the right explanation and his chances of amending this mistake seemed slim. We are also not told by O'Connor how reasons can influence us without causing the action. Thus, with Clarke, Van Miltenburg, and Feldman and Buckareff, I judge O'Connor's theory on reasons to be insufficient.

We also saw some negative points in chapter two however. Though his opponents failed to show that O'Connor is incoherent, they did point to other difficulties. How can O'Connor explain that an action happens at some exact moment? How can the agent make her final decision? At the final decision, only the agent herself and nothing about her determines the outcome. On what ground does she choose? Of course, O'Connor has indicated that he cannot explain some of these things because agent-causal powers is a primitive notion. Saying that his theory is a complete failure because it lacks certain explanations is therefore not fair. Still, the missing explanations do leave us at least partially unsatisfied. This should be addressed, if possible.

Chapter five was atypical in the sense that I did not directly criticized O'Connor or his opponents there. I shortly touched on some issues O'Connor has to deal with concerning empirical data and current scientific research. But more importantly, I noted that his claims about the phenomenology of action were underdeveloped and deserve more serious and systematic research. This was especially so because O'Connor thinks that phenomenology can be used to justify, to some extent, his attempts at formulating an agent-causal account of action and freedom. I suggested that systematic research should be done to either corroborate or reject O'Connor's ideas concerning our experiences.

Ultimately, I conclude that O'Connor's theory is not problematic to the point of being impossible or useless. It is not incoherent and it manages to give us an intelligible account of what human freedom is like. Still, I think that the theory has some significant flaws that should be addressed. There is a number of things that O'Connor cannot explain. His failure to present us with a suitable account of the influence of reasons on our actions is also significant. O'Connor could and does argue that his theory can still be considered valuable without a theory on reasons[69], but it still seems a considerable disappointment. All in all, I think that O'Connor has presented us with a theory that is possible and intelligible, but not highly desirable. It might be better to look for alternatives. An event-causal theory of human action might be a possible candidate, since O'Connor admits that events can cause actions. After all, he considers certain unfree actions to have an event-causal structure. Of course, the limits of this thesis prevent us from investigating whether such optimism is justified. It is also possible that all alternative theories of free action fail or have even worse vices. In that case, it would make perfect sense to adopt O'Connor's theory, since it at least is coherent and also has a number of

[69] O'Connor, *Persons and Causes,* 100.

adequate explanations. A serious investigation into alternative theories of free action, and a systematic investigation of the phenomenology of action, would be a good way to build on the work done here.

Literature

Clarke, Randolph. *Libertarian Accounts of Free Will*. Oxford: Oxford University Press, 2003.

Davidson, Donald. ''Actions, Reasons and Causes.'' *The Journal of Philosophy* 60 (1963), 685-700.

Feldman, Richard, and Andrei A. Buckareff. ''Reasons Explanations and Pure Agency.'' *Philosophical Studies* 112 (2003), 135-145.

Ginet, Carl. *On Action.* Cambridge: Cambridge University Press, 1990.

Mele, Alfred, ''Ultimate Responsibility and Dumb Luck.'' *Social Philosophy and Policy* 16 (1999), 274-293.

Nahmias, Eddy, Stephen Morris, Thomas Nadelhoffer and Jason Turner. ''The Phenomenology of Free Will.'' *Journal of Consciousness Studies* 11 (2004), 162-179.

Nahmias, Eddy and Jason Turner. ''Are the Folk Agent-Causationists?'' *Mind & Language* 21 (2006), 597-609.

O'Connor, Timothy. *Persons and Causes.* Oxford: Oxford University Press, 2000.

O'Connor, Timothy. ''Agent-Causal Theories of Freedom.'' In *The Oxford Handbook of Free Will (Second Edition)*, edited by Robert Kane, 309-328. Oxford: Oxford University Press, 2011.

Pereboom, Derk. *Living Without Free Will.* Cambridge: Cambridge University Press, 2001.

Pereboom, Derk. *Free Will, Agency and Meaning in Life.* Oxford: Oxford University Press, 2014.

Strawson, Galen. *Freedom and Belief*. Oxford: Clarendon Press, 1986.

Strawson, Galen. ''The Bounds of Freedom.'' In *The Oxford Handbook of Free Will (Second Edition)*, edited by Robert Kane, 441-460. Oxford: Oxford University Press, 2011.

Van Inwagen, Peter. ''Free Will Remains a Mystery.'' *Philosophical Perspectives* 14 (2000), 1-14.

Van Miltenburg, Niels. *Freedom in Action.* Utrecht: Quaestiones Infinitae, 2015.