

Atle Ottesen Søvik

Free Will, Causality and the Self

Philosophische Analyse/ Philosophical Analysis



Herausgegeben von/Edited by
Herbert Hochberg, Rafael Hüntelmann,
Christian Kanzian, Richard Schantz, Erwin Tegtmeier

Band/Volume 71

Atle Ottesen Sørvik

Free Will, Causality and the Self

DE GRUYTER

ISBN 978-3-11-047431-2
e-ISBN (PDF) 978-3-11-047468-8
e-ISBN (EPUB) 978-3-11-047446-6
ISSN 2198-2066

Library of Congress Cataloging-in-Publication Data

A CIP catalog record for this book has been applied for at the Library of Congress.

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at <http://dnb.dnb.de>.

© 2016 Walter de Gruyter GmbH, Berlin/Boston
Printing: CPI books GmbH, Leck
♻️ Printed on acid-free paper
Printed in Germany

www.degruyter.com

Foreword

This book took some years to write. The discussion on free will moves fast forward and I have been helped by many people in getting to know and think through details:

Many thanks to the following people for reading the whole manuscript at one time or another and offering many valuable comments: Robert Kane, Bob Doyle, Ragnar Mogård Bergem, Lars Fredrik Svendsen, Helen Steward, Olav Gjelsvik, Bjørn Ramberg, Michael Gisinger, and two anonymous reviewers.

Thanks to the following people for reading parts of the manuscript and offering valuable comments: Jonathan Schaffer, Oskar Skarsaune and Kristin Graff-Kallevåg.

I am grateful for brief conversations on the matter with the following people: Randolph Clarke, Alfred Mele, Michael McKenna, Chandra Sripada. Also, thanks to many of the people at CSMN/IFIKK at the University of Oslo for discussions over coffee; Einar Duenger Bøhn, Andreas Brekke Carlsson, Andreas Strand, Edmund Henden, Monica Roland and Hedda Hassel Mørch.

Thanks to the people involved at DeGruyter: Rafael Hüntelmann, Olena Gainulina and Maik Bierwirth, and to Kjetil Borgenvik at MF for great technical help with the manuscript.

Finally, I am grateful to my family who would have made my life a great joy even if free will had not existed: Andreas, Jenny, Kristian and Elise.

Oslo, February 2016.

Contents

1 Introduction — 1

- 1.1 The Problem of Free Will—Main Positions and Problems — **1**
- 1.2 The Theory Proposed in this Book — **10**
- 1.3 Methodological Reflections — **17**

2 Causality — 21

- 2.1 Part 1: Jonathan Schaffer's Understanding of Causality — **21**
 - 2.1.1 To what category do causes and effects belong? — **22**
 - 2.1.2 How many relata are there in the causal relation? — **22**
 - 2.1.3 How are causes and effects connected? — **23**
 - 2.1.4 How are causes selected from other conditions? — **28**
- 2.2 Part 2: Evaluation of Schaffer's Understanding of Causality — **29**
 - 2.2.1 To what category do causes and effects belong? — **31**
 - 2.2.2 How many relata are there in the causal relation? — **33**
 - 2.2.3 How are causes and effects connected? — **34**
 - 2.2.4 How are causes selected from other conditions? — **49**

3 The Self — 56

- 3.1 Biological Background — **59**
- 3.2 Brain and Mind — **62**
- 3.3 Emotion — **72**
- 3.4 Memory — **76**
- 3.5 Consciousness — **78**
- 3.6 The Self — **81**
- 3.7 The Function of Consciousness — **89**
- 3.8 Thinking — **95**
- 3.9 Desire — **97**

4 Free Will — 106

- 4.1 What Is Free Will? — **106**
- 4.2 Alternative Possibilities — **114**
- 4.3 Determinism — **116**
- 4.4 Independence of the Self — **121**
- 4.5 Responsibility — **126**
- 4.6 Free Will and Responsibility in Various Cases — **139**

VII — Contents

5 Answers to Objections — 157

5.1 Main problems for Libertarians — **157**

5.2 Main Problems for Compatibilists — **162**

5.3 Other Problems — **163**

5.4 Conclusion — **170**

Bibliography — 173

Name index — 179

Subject index — 181

1 Introduction

In this chapter I first present the main positions in the free will debate—compatibilism and libertarianism. I also explore their subtypes: for compatibilism, mesh theories and reasons-responsive theories; for libertarianism, non-causal, agent-causal and event-causal theories. I then present the main problems these positions each face. For compatibilism, this includes the consequence argument, the manipulation argument and the zygote argument. Problems for libertarianism include the luck problem, the assimilation argument and the disappearing agent problem. After this, I present the theory I will propose in this book, indicating how it deals with all of these problems. I end with some methodological reflections.

1.1 The Problem of Free Will—Main Positions and Problems

There are three main questions that a theory of free will should answer: What does “free will” mean? Do we have free will? Is the proposed theory of free will coherent, given that free will seems incompatible with both determinism and indeterminism?

What does the term “free will” refer to? There are many different definitions of free will, and one should ask which understanding of free will is being considered, whether it is being affirmed or rejected. It may well be that a strong form of free will is rejected while a weaker form is affirmed. In ordinary language, a minimum requirement for what it means to have free will is to say that it is the freedom persons must have in order for it to be meaningful to hold them responsible for their actions. (Like free will, though, responsibility is understood in many different ways.) A common definition of free will is to say that (1) it is “up to us” what we choose between several alternatives, and (2) the source of the choice is in us, and not outside of us or in something else that we cannot control.¹ Even if this description does not apply to every choice, most people will say that they experience such free will at least in some of their choices. Still, both parts of even this definition are contested and can be further defined in various ways: What does it mean to be the source of a choice? Are alternative possibilities necessary for free will? How should such alternative possibilities be

¹ Robert Kane, *The Oxford Handbook of Free Will*, 2nd ed. (Oxford: Oxford University Press, 2011), 5.

understood? Some philosophers will say that you have no free will, others will say that you have free will in a weak sense of the term, and still others will say that you have free will in a strong sense of the term. It is my goal in this book to try to find out how free your will really is.

Among those who affirm free will, compatibilists and libertarians disagree on whether free will is compatible with the idea that all events are determined. This idea, *determinism*, can be defined in many ways.² One common understanding refers to the belief that everything that happens must necessarily happen exactly as it does because of previous causes; therefore, at any one time, there is only one possible future.³ Even within this definition there are many varieties, since destiny, God's predestination, logical necessities, or physical causes can be that which determines the outcome. In the free will debate today, the focus is usually on physical determinism, which is the view that previous physical causes plus the laws of nature determine one future with physical necessity.⁴ At any point of time, the rest of history is then implied by the state of the world at that time. Such physical determinism will be the focus in this book as well, since I believe that is the strongest and most common challenge to the question of free will.⁵ This is a metaphysical position, and this is what I mean by "determinism" in this book.

In the discussion on determinism, Compatibilists believe that determinism is compatible with having free will. For a long time the most common critique of

² According to J. H. Sobel, there are at least 90 different understandings of the term "determinism." See Bishop, in *ibid.*, 84.

³ Note that determinism does not necessarily mean predictability, since a system can be determined and at the same time chaotic and/or incomputable. Determinism is an ontological question; predictability is an epistemological question (Nicholas Saunders, *Divine Action and Modern Science* (Cambridge: Cambridge University Press, 2002), 87–90, 177–78). "Only one possible future" means that only one specific series of events is already determined to be the content of the future.

⁴ The term "physical" is notoriously difficult to define, but a rough definition that should suffice in this context is that "physical" is that which is to some degree measurable and quantifiable.

⁵ One definition is as follows: "The state of the universe at any time is wholly and unequivocally determined by the state of the universe at prior times, and the physical laws of nature." (Hodgson, in Kane, *The Oxford Handbook of Free Will* (2nd ed.), 57). This kind of physical determinism may be further divided into causal physical determinism and block universe physical determinism. My reference is to the causal version (See Bishop, in *The Oxford Handbook of Free Will* (Oxford: Oxford University Press, 2002), 113–14). We shall return to the question of how to understand causality in chapter two—for now, suffice it to say that physical causes are immanent events in the world.

compatibilism was the consequence argument. Roughly, this argument is that if determinism is true, then what happens in the future is determined by laws of nature and events that took place previously, even in the distant past. Even before any humans existed it was determined what the content of the future would be for all humans. Since the future was thus determined before our birth, it cannot be up to us what happens among different alternatives, and thus we cannot have free will.⁶

Despite this widely debated argument, compatibilists believe determinism is compatible with having free will. Compatibilists today will usually say that it does not matter that only one specific future is physically possible. Rather, they will focus instead on what the inner mental life of an agent must be like in order for the agent to be free. One strand of contemporary compatibilism is the so-called *mesh theories*, which hold that a person is free when she has the right connections or “mesh” between internal parts of her mental life.⁷ Another strand is *reasons-responsive theories*. According to these theories a person is free when her actions are based on a rational response to reasons for action.⁸ These different compatibilist understandings of free will do not require indeterminism, and so free will is argued to be compatible with determinism and in no need of alternative possibilities.

However, an argument other than the Consequence Argument has been in focus lately against compatibilism, and that is Derk Pereboom’s four-case manipulation argument. This argument presents four cases, from a clear manipulation case to a deterministic world, where the point is to show that there are no relevant differences between the cases. Since the first case clearly seems to be a

⁶ Peter Van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983).

⁷ For example, there are hierarchical mesh theories, such as Harry Frankfurt’s theory. Frankfurt argues that we have several desires whose object is an action or a state, which he calls first-order desires. But we also have desires whose object is a first-order desire, and these he calls second-order desires. The second-order desires are internal responses to the first-order desires, which one may like or dislike. According to Frankfurt, we are free when our second-order desires approve our first-order desires, because only then do we have the will we want. (Harry Frankfurt, “Freedom of the Will and the Concept of a Person,” *Journal of Philosophy* 68(1971)).

⁸ See for example Susan R. Wolf, *Freedom within Reason* (New York: Oxford University Press, 1990), or John Martin Fischer and Mark Ravizza, *Responsibility and Control: A Theory of Moral Responsibility*, Cambridge Studies in Philosophy and Law (Cambridge: Cambridge University Press, 1998).

case of no free will, the charge is to explain the relevant difference between case 1 and case 4.⁹

⁹ Here are the four cases, quoted from Derk Pereboom, *Free Will, Agency, and Meaning in Life* (New York: Oxford University press, 2014), 76–79:

Case 1: A team of neuroscientists has the ability to manipulate Plum's neural states at any time by radio-like technology. In this particular case, they do so by pressing a button just before he begins to reason about his situation, which they know will produce in him a neural state that realizes a strongly egoistic reasoning process, which the neuroscientists know will deterministically result in his decision to kill White. Plum would not have killed White had the neuroscientists not intervened, since his reasoning would then not have been sufficiently egoistic to produce this decision. But at the same time, Plum's effective first-order desire to kill White conforms to his second-order desires. In addition, his process of deliberation from which the decision results is reasons-responsive; in particular, this type of process would have resulted in Plum's refraining from deciding to kill White in certain situations in which his reasons were different. His reasoning is consistent with his character because it is frequently egoistic and sometimes strongly so. Still, it is not in general exclusively egoistic, because he sometimes successfully regulates his behavior by moral reasons, especially when the egoistic reasons are relatively weak. Plum is also not constrained to act as he does, for he does not act because of an irresistible desire—the neuroscientists do not induce a desire of this sort.

Case 2: Plum is just like an ordinary human being, except that a team of neuroscientists programmed him at the beginning of his life so that his reasoning is often but not always egoistic (as in Case 1), and at times strongly so, with the intended consequence that in his current circumstances he is causally determined to engage in the egoistic reasons-responsive process of deliberation and to have the set of first- and second-order desires that result in his decision to kill White. Plum has the general ability to regulate his actions by moral reasons, but in his circumstances, due to the strongly egoistic nature of his deliberative reasoning, he is causally determined to make the decision to kill. Yet he does not decide as he does because of an irresistible desire. The neural realization of his reasoning process and of his decision is exactly the same as it is in Case 1 (although their causal histories are different).

Case 3: Plum is an ordinary human being except that the training practices of his community causally determined the nature of his deliberative reasoning processes so that they are frequently but not exclusively rationally egoistic (the resulting nature of his deliberative reasoning processes are exactly as they are in Cases 1 and 2). This training was completed before he developed the ability to prevent or alter these practices. Due to the aspect of his character produced by this training, in his present circumstances he is causally determined to engage in the strongly egoistic reasons-responsive process of deliberation that issue in his decision to kill White. While Plum does have the general ability to regulate his behavior with moral reasoning, in virtue of this aspect of his character and his circumstances he is causally determined to make his immoral decision, although he does not decide as he does due to an irresistible desire. The neural realization of his deliberative reasoning process and of the decision is just as it is in Cases 1 and 2.

Case 4: Everything that happens in our universe is causally determined by virtue of its past states together with the laws of nature. Plum is an ordinary human being, raised in normal circumstances, and again his reasoning processes are frequently but not exclusively egoistic,

Alfred Mele has a similar argument called the zygote argument: imagine a goddess creating a zygote at exactly the right time and place with the exact right structure in a deterministic universe. She does this because she knows that the zygote will then become a man (Ernie) who at an exact point of time will do something the goddess wants done—for example, kill his grandmother. Ernie will be, like any other person in a deterministic universe, considered by compatibilists to be free, but many will have the intuition that he was not responsible for killing his grandmother since the goddess had planned things so that this had to happen. Yet, since he is like any other person in our world if the world is determined, it seems that if he is not responsible, no one else is either.¹⁰

Even if one disagrees over how strong the manipulation argument and the zygote argument is against compatibilism, I think there can be little doubt that, if the future is already determined before we are born, we do not have a strong form of free will. It is not up to us to change the future into anything else than what was already determined before we were born. Libertarians, on the other hand, think that we do have a stronger form of free will than this. They hold that we can be the source of our choices in a more fundamental sense than what compatibilists will allow, but that requires an indeterministic world where different futures are possible and where it is up to us to influence what the future will be like.

There are three main positions among libertarians distinguished by how they understand the causality involved in free choices. Non-causalists believe that free actions are not caused at all, but are intelligible in the light of the purpose of the action. Agent causalists believe that there is a unique and irreducible kind of causation that only free agents can employ. Event causalists deny that actions have special causes, but believe instead that all causes are of the same kind: they think that events cause events, both in the mind and in the world in general.

Those who defend the strongest form of free will are the non-causalists and the agent causalists. Non-causalists argue that human action should be explained by intentions or reasons instead of causes, and that these are not reduc-

and sometimes strongly so (as in Cases 1–3). His decision to kill White issues from his strongly egoistic but reasons-responsive process of deliberation, and he has specified first- and second-order desire. The neural realization of Plum's reasoning process and decision is exactly as it is in Cases 1–3; he has the general ability to grasp, apply and regulate his actions with moral reasoning, and it is not because of an irresistible desire that he decides to kill.

10 Alfred R. Mele, *Free Will and Luck* (New York: Oxford University Press, 2006), 188–89.

ible to ordinary event causes.¹¹ A classic charge against this view was leveled by Donald Davidson.¹² He pointed out that even if a person has a reason for doing something, that does not mean that his reason is what actually caused the event to happen. People often experience having competing reasons for doing different things when they act. The challenge to non-causalists is to explain what links the personal reason to the action. Agent causalists like Timothy O'Connor hold that agents are enduring irreducible substances who have a unique ability to perform actions.¹³ Agent causalists are typically criticized for appealing to both a mysterious agent and a mysterious form of causation, which does not fit into the ordinary scientific world view.¹⁴ Nor do they explain how reasons make actions happen, for in virtue of what does the agent control her actions? Agent-causation can be argued to be an irreducible phenomenon,¹⁵ but I shall argue later that it is superfluous to add anything to normal event-causation. This will be my main argument against non-causal and agent-causal libertarian theories—that our behavior can be explained sufficiently in event-causal terms so that there is no good reason to believe in extra agency or causation beyond that.

In my view, the most plausible of the libertarian theories are the event-causal theories. Event-causalists hold that mental events can cause free actions. There are two main event-causal theories, distinguished by where in the deliberation process they locate indeterminism.¹⁶ The advocates of *centered* event-causal theories believe that there is indeterminism until and in the moment of choice, whereas advocates of the *deliberative* event-causal theories hold that there is indeterminism early in the deliberation process, creating different ideas in the mind (alternative possibilities), but that the rest of the deliberation process is determined.¹⁷

11 An example of such a theory can be found in Carl Ginet, *On Action*, Cambridge Studies in Philosophy (Cambridge: Cambridge University Press, 1990).

12 Donald Davidson, "Actions, Reasons and Causes," *The Journal of Philosophy* 60, no. 23 (1963).

13 O'Connor, in Kane, *The Oxford Handbook of Free Will* (2nd ed.), 309–28.

14 Pereboom, *Free Will, Agency, and Meaning in Life*, 65–69.

15 For example, E. J. Lowe argues that both causality and agent-causality are irreducible concepts. See E. J. Lowe, *A Survey of Metaphysics* (Oxford: Oxford University Press, 2002), chapters eight to eleven.

16 Indeterminism simply means that more than one future is possible, so when I and others I refer to speak about the location of indeterminism, the point is to speak about the source of indeterminism; where do the indeterministic effects arise?

17 The distinction between centered and deliberative event-causal theories is from Randolph K. Clarke, *Libertarian Accounts of Free Will* (New York: Oxford University Press, 2003), 57, 71.

Since event-causal libertarian theories are close to my own theory, I shall spend a little time in presenting them here. I start with the centered event-causal libertarian theory of Robert Kane.¹⁸ Kane thinks that free will is fundamentally about the ultimate source of action being in us.¹⁹ More precisely, the requirement is that “*To be ultimately responsible for an action, an agent must be responsible for anything that is a sufficient cause or motive for the action’s occurring.*”²⁰ This means that the requirement of alternative possibilities is not necessary for all our actions. But it is necessary in some specific early choices in which we formed our own characters, according to Kane. He calls such actions self-forming actions, or SFAs. Even if one could not have done otherwise in some situations, one is still responsible if the reason that one could not do otherwise was earlier SFAs. For example, if you have formed your character through SFAs so that it is now impossible for you to lie, you are still free, responsible and praiseworthy for not lying in situations where you could.²¹ As long as we are free to make some SFAs, we can be free, but if the world is determined, then none of our actions are SFAs and then we are not free.

But this seems to result in an infinite regress, for would not those earlier choices depend on even earlier choices, and so on indefinitely?²² Kane’s response to this criticism is that the regress is ended if there is an action in the agent’s past that lacked sufficient motive.²³ There could be a situation in which the agent did not know what to do and so did not set her will before the action occurred. Then the action would set the will in the very act of choosing; Kane calls such actions “will-setting actions” (which are the same as self-forming actions). He adds that in order for such actions to provide us with free will, they must have been such that the agent could act voluntarily, intentionally and rationally in more than one way when she acted. If the action happened as an accident, it would not have made the agent the ultimate source; rather, the accident would be the source. But if the agent had a motive for both alterna-

18 For a detailed presentation of Kane’s theory and a critique of it, see Atle O. Søvik, “A Critique of Robert Kane’s Theory of Free Will,” (Under review).

19 Kane, “Libertarianism,” John Martin Fischer, *Four Views on Free Will*, Great Debates in Philosophy (Oxford: Blackwell Pub., 2007), 13–14.

20 Kane, “Libertarianism,” *ibid.*, 14. Emphasis in original text.

21 “Could” must here be understood in the sense that it was type physically possible.

22 Kane, in Fischer, *Four Views on Free Will*, 19–20.

23 According to Kane, we have a “sufficient motive” for doing something when our will is set one way on doing so before and when we act (Kane, *ibid.*, 19).

tives, then she is the ultimate source of the choice no matter what she chooses, so the regress stops there.²⁴

Even if such a choice is undetermined, Kane still thinks it can be a rationally willed choice. To argue this, he offers the example of a businesswoman on her way to an important meeting who witnesses an assault. In this situation she has reasons to stop and reasons to move on, and she does not know what to do. The conflicting motives stir up a chaos in the brain, which is sensitive to undetermined events at the micro level of quantum mechanics. In this situation the woman must make an effort to choose and no matter what she chooses, it will be for a reason. When she decides, that decision sets her will.²⁵

The most common critique of Kane's theory is that it runs into a problem of luck. Let us say there is a 70% probability that Jack will decide to have pancakes for breakfast. If history were rolled back a hundred times and played again up to the moment of choice, Jack would decide to have pancakes 70 times and something else 30 times. But if the exact same history up to the moment of choice can give completely different choices—which Robert Kane argues it can²⁶—it seems to be a matter of luck as to what Jack decides to do. The same point can be made with identical worlds up to the moment of choice, where Jack1 and Jack2 make different choices.²⁷

The luck argument has been strengthened by the assimilation argument. Seth Shabo has argued a case where an indeterministic device in a person's brain determines her choices but sometimes it does not work, and then compared these situations in a rollback of histories. The argument then is that there is no relevant difference between a rollback of histories where the indeterministic device works or not—in any case there will be an undetermined outcome where the person sometimes chooses A and sometimes chooses B.²⁸

Kane's theory is a *centered* event-causal theory since it locates indeterminism in the moment of choice. Deliberative event-causal theories try to reduce the luck component by locating indeterminism at an earlier point in the deliberation process. Such models are also called two-stage models, since the delibera-

²⁴ Kane, *ibid.*, 20.

²⁵ Kane, *ibid.*, 26–28.

²⁶ Kane, *ibid.*, 23.

²⁷ Randolph Clarke distinguishes between the rollback argument with a rollback of history, and the luck-objection as the inter-world case (Clarke, in Kane, *The Oxford Handbook of Free Will* (2nd ed.), 323–24).

²⁸ Seth Shabo, "Why Free Will Remains a Mystery," *Pacific Philosophical Quarterly* 92, no. 1 (2011), "Assimilations and Rollbacks: Two Arguments against Libertarianism Defended," *Philosophia* 42, no. 1 (2014).

tion process comprises two stages. First there is an indeterministic stage in which alternatives for actions are generated in the mind. This is followed by a deterministic stage in which one alternative for action is selected.

One of the best such proposals is Alfred Mele's *daring soft libertarianism*. Mele argues that whereas the other models shun luck and only include indeterminism to avoid determinism, this model embraces luck while still maintaining that the agent can be in control.²⁹ The point is that the deliberation process is indeterministic, so that it is partly a matter of luck what the agents end up choosing. But the agent learns from experiences over time and the agent's evaluations of these experiences influences how likely it is that the same choice will be made later. In this way, the agent shapes her own character over time and this also explains why we hold children less responsible than adults for what they do.³⁰

Neil Levy argues that libertarians can try to reduce the luck component by making their theories almost compatibilist.³¹ But he does not think that Mele's strategy with including luck in the history of an agent works, since luck has been a part of every choice, and you cannot solve the problem of luck with adding more luck.³² Even compatibilists have a luck problem, according to Levy, since it is also a matter of luck what such agents come to think about or desire.³³

In addition to the luck problem, the other main argument against event-causal theories is the problem of the disappearing agent. It seems that on event-causal approaches, choices reduce to desires, beliefs and bodily movement, and the agent disappears.³⁴ If everything is just natural causal processes occurring, where is the free agent?

As seen, the different kinds of compatibilisms and libertarianisms run into different problems. This has led various philosophers to conclude that we do not have free will, since free will is incompatible with both determinism and indeterminism.³⁵ But most philosophers still hold on to the idea that we have free

²⁹ Mele, *Free Will and Luck*, 117.

³⁰ *Ibid.*, 122–23, 31–32.

³¹ Neil Levy, *Hard Luck: How Luck Undermines Free Will and Moral Responsibility* (New York: Oxford University Press, 2011), 77.

³² *Ibid.*, 89.

³³ *Ibid.*, 90.

³⁴ This is mentioned as a main objection against event-causal theories in, for example, Pereboom, *Free Will, Agency, and Meaning in Life*, 31–33, and Helen Steward, *A Metaphysics for Freedom* (Oxford: Oxford University Press, 2012), 62.

³⁵ Pereboom, *Free Will, Agency, and Meaning in Life*, 3.

will, since this seems best to fit our experience of having free will and being responsible for our actions. Now that we have a conceptual map of the main positions and main problems, I am ready to locate my own theory as a newcomer to this map and indicate how I will relate to the different problems.

1.2 The Theory Proposed in this Book

If the compatibilists are right, we have free will only in a very limited sense since our whole future was determined before we were born. The manipulation argument and the zygote argument illustrate this well. Luck is also a real limitation to the degree of control that we can have, and I cannot see that there are good arguments in favor of non-causal or agent-causal libertarian theories. I do think, though, that we have a stronger form of free will than that which compatibilists hold, while I do not think the problem of luck is as big as it is often claimed to be.

The theory I propose in this book is a new way of navigating between compatibilism and libertarianism, and it is closest to event-causal libertarianism. It is libertarianism, since I believe that the world is indeterminated at the macro level of humans, and that this is required for free will in a strong sense. I reject determinism and compatibilism, and think that we can be fundamental sources of our own choices. However, event-causal libertarianism always locates the indeterminism inside the mind of the agent. I, however, claim that there is no need for such agent-internal indeterminism. It suffices that, for free will in a strong sense, there is some indeterminism somewhere in the world making more than one future physically possible. External indeterminism is all that is required to have free will in a strong sense. Removing the internal indeterminism reduces the luck component, while still making more than one future possible, thereby avoiding the manipulation and zygote arguments.

It has been claimed that it is preposterous to base a theory of free will on external indeterminism,³⁶ but I shall try anyway. I will also respond to this criticism in my presentation. First, in order to show how it is possible to be the fundamental source of our choices on such conditions, it is necessary to give a detailed understanding of causation and the self. Such understandings will be presented in chapters two and three. This will clarify what it means to be an agent and what it means to be the source of a choice. I strive to give detailed

³⁶ Alfred R. Mele, *Autonomous Agents: From Self-Control to Autonomy* (New York: Oxford University Press, 1995), 195–96.

empirical content to such concepts as *agent*, *self*, *choice*, *cause*, *source*, and *control*, and will show that many problems can be solved when these are given clear empirical content.

In detail, I will structure my case as follows:

Chapter two is my understanding of causation. I argue that an agent can be the cause or ultimate source of her choices even if everything that happens in her mind is a causal process where only one thing may be token physically possible to do. In order to argue this, I must present a detailed and empirical theory about causation. I must also explain why causation should be thought of as contrastive and that we select contrasts and causes based on interests and expectations. This allows me to be clear about what it is to be the ultimate cause or source of one's actions, and to understand how agents can cause actions this way.

In chapter three, I present my theory of the self, which is informed to a large degree by the neuroscientist Antonio Damasio. His theory of the self lets me explain how a person's self can be the cause of her actions. I discuss in some detail how the mind works, and focus on giving empirical content to the concepts, trying to relate frameworks with a first-person and a third-person perspective. The work done here should be of interest to those who think about free will even if they disagree with everything else I say before that. Damasio distinguishes between the proto-self, the core self, and the autobiographical self. Especially important for me is the understanding of our autobiographical self, and how it can change the desires we feel, so that the autobiographical self then becomes the cause of the desires that cause our actions. Since the world is indetermined and causation contrastive, I shall argue that there will be cases where the autobiographical self is the ultimate cause of a person's action, and that it is the ultimate cause of its own content.

Damasio's work lets us describe a deliberation process as a causal process, where a situation triggers alternatives for action that a person desires with different degrees of strength, and activated memories from the autobiographical self mark the alternatives with feelings which change the strength of the desire for each alternative. When the strongest desire reaches a threshold for action, motor neurons are activated and the action is carried through.

With this detailed theory of the self and deliberation as background, I present my theory of free will in chapter four. I start by describing many different kinds of deliberation processes wherein an agent (which in my terminology is the same as a person, which I define as a human body with a mind and a core self) and her autobiographical self can be involved to varying degrees. I argue

that a person has free will if (but not only if) her autobiographical self changes her initial desire into a different desire, which causes the choice that is made, because the person is then the cause of her choice between alternatives, and the source of the choice was within her.³⁷

Concerning the requirement of alternative possibilities, in any situation there may be only one thing that it was *token* physically possible for a person to do. But as long as the person imagines different alternatives, the person can be rightly selected as the cause of the desire that causes the choice that is made—all the imagined alternatives need not be *token* physically possible for the person to do as long as they are *type* physically possible.³⁸ This presupposes that the world cannot be determined if we are to be free, for then our actions would have been determined before we were born, and it would not be right to select the autobiographical self as the cause of the choice between A and B. But if there is some indeterminism in the world, with undetermined effects at the macro level of humans, then it will often be right to select persons as the causes of the choices they make, or so I shall argue in chapter four on the basis of the discussion of causality in chapter two.

People can have more or less independent autobiographical selves, depending on how much they have changed their autobiographical selves from within

37 “Choice” is an ambiguous term (Nicholas Rescher, *Free Will: A Philosophical Reappraisal* (New Brunswick, NJ: Transaction Publishers, 2009), xii). Before I choose I *have a choice*, say between alternatives A and B. A and B are possible *choices* (alternatives), and I can *make a choice* between them. When the deliberation process is over, I have *made a choice*, which means I have decided to do either A or B. There are several stages from deliberation to action. First, there is the deliberation process wherein different alternatives and desires are considered, then one of the alternatives is selected as the one to act upon, and finally an alternative is acted upon. The alternative that is acted upon is usually the one that was selected, but if there is sufficient time between the selection and action, new deliberation may take place, and a change of mind can occur. “Choice” and “decision” are ambiguous terms, but I try to clarify what I mean, and usually I mean by “choice” or “decision” the selection of the alternative that was finally acted upon. Alfred Mele suggests “intention” for the selection of the alternative to act upon, and argues that one can intend to do something without having decided it (Alfred R. Mele, *Springs of Action: Understanding Intentional Behavior* (New York: Oxford University Press, 1992), 141). It is true that “choice” suggests that several alternatives have been considered, whereas if a person only has one desire and only considers one alternative, they can act intentionally without having made a choice. All these parts of the deliberation process will be presented in detail in chapter three. I shall focus in this book on decisions about how to act, not cases where one decides what one believes to be true.

38 Something may be *type* physically possible if it is physically possible in general. But maybe the same thing is *token* physically impossible in a certain situation, and token physically possible in other specific situations.