



© Jason Douglas Runyan 2014

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted his right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2014 by
PALGRAVE MACMILLAN

Palgrave Macmillan in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of Houndmills, Basingstoke, Hampshire RG21 6XS.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave Macmillan is the global academic imprint of the above companies and has companies and representatives throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–1–137–32948–6

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

A catalogue record for this book is available from the British Library.

A catalog record for this book is available from the Library of Congress.

Typeset by MPS Limited, Chennai, India.

Contents

<i>List of Figures</i>	ix
<i>Acknowledgements</i>	x
1 Introduction	1
1.1 What is at issue	1
1.2 An outline of the work ahead	6
1.3 A different approach to a familiar topic: Undercutting the contemporary compatibilist/ incompatibilist debate	9
2 Libet-Style Experiments and Volitions	16
2.1 Introduction	16
2.2 Libet-style experiments and the problem they present for volitional accounts	18
2.3 Volitional rebuttals to Libet-style problems	23
2.4 Conceptual and empirical grounds for rejecting volitional accounts	30
2.5 Conclusions	42
3 The Need for an Analysis of Human Agency	45
3.1 Introduction	45
3.2 Volitional accounts, action and awareness	46
3.3 An analytic approach reaffirming a broadly-Aristotelian account of human agency and voluntary conduct	53
3.4 Conclusions	55
4 An Aristotelian Account of Human Agency	57
4.1 Introduction	57
4.2 Wants and choices	59
4.3 Human agency	61
4.4 Awareness and choosing	72
4.5 Conclusions	83
5 Compatibilist Concerns	86
5.1 Introduction	86
5.2 Incompatibilist implications	88
5.3 Possible compatibilist objections and incompatibilist replies	96

5.4	Implications for compatibilist accounts of deliberation	98
5.5	Conclusions	101
6	Choices and Voluntary Conduct	103
6.1	Introduction	103
6.2	Bodily motion, action and choices	105
6.3	Power to choose and voluntary agency	106
6.4	Frankfurt-type examples	111
6.5	The illusion of voluntary agency and social psychology	115
6.6	Conclusion	124
7	Neuronal Mechanisms and Voluntary Agency	126
7.1	Introduction	126
7.2	Human agency and subpersonal causes	134
7.3	The distinctiveness of complex neural systems	144
7.4	Why upholding exhaustively mechanistic theories is contradictory or self-refuting	155
7.5	Conclusions	165
8	A Metaphysical Framework: Voluntary Agency, Emergence and Downward Causation	166
8.1	Emergence and voluntary agency	166
8.2	Toward an emergentist account of voluntary agency	174
8.3	Concluding remarks: Two contrasting depictions of the universe	180
	<i>Notes</i>	193
	<i>Index</i>	228

1

Introduction

1.1 What is at issue

The LORD said to Cain, ‘Why are you angry? And why has your countenance fallen? If you do well, will you not be accepted? And if you do not do well, sin is lurking at the door; its desire is for you, but you must master it.’

Cain said to his brother Abel, ‘Let us go out to the field.’ And when they were in the field, Cain rose up against his brother Abel, and killed him. Then the LORD said to Cain, ‘Where is Abel your brother?’ He said, ‘I do not know; am I my brother’s keeper?’ And the LORD said, ‘What have you done? Listen; your brother’s blood is crying to me from the ground!’¹

In this ancient murder story, by all accounts, if any action is voluntary the killing of Abel is. It is, thus, an act for which Cain is morally responsible. Ancient audiences had no problems recognising this, and neither do we. However, over, at least, the past 350 years or so there have been two divergent accounts of what distinguishes conduct—like the killing of Abel—as voluntary.

According to the first account—which I will call the *Aristotelian account* since its roots can be traced back to views held by Aristotle—a voluntary action is, as a first approximation, the exercise of a two-way power. That is, it is an act one performs when one is able to act and able to not act on that occasion.² On such occasions, it is up to the individual whether or not they act. According to this way of thinking, the killing of Abel is voluntary, roughly, because Cain brought about Abel’s death, knew what he was doing (e.g., that he was not doing ‘well’,

2 *Human Agency and Neural Causes*

giving in to sinful desire, ending Abel's life), knew he could have not acted as he did (e.g., that he could act 'well', 'master' sinful desires, not kill Abel), and was not under duress.

According to the second account—which I will call the *volitional account*—a voluntary action is, roughly, bodily motion that has a certain mental event in its causal history.³ In one version—what is called an *event-causal* account—bodily motion is caused by a certain kind of mental event of which one is occurrently conscious (or conscious of in the way one is conscious of a sensation one is feeling). In another version—what is called an *agent-causal* account—bodily motion is caused by the individual when they perform a certain mental act; the performance of which is a mental event. In either version, the pivotal mental event has been called by various names, including a 'conscious volition', 'conscious choice', 'conscious intention' and 'mental act of will'. According to volitional ways of thinking, in general terms, the killing of Abel is voluntary: granted Cain's body moved in a way that led to Abel's death, and that this motion was caused by either a particular mental event, or by Cain when he performed a certain mental act.

So according to an Aristotelian account, voluntarily acting is, roughly, a way of exercising a two-way power by one who is informed of certain things and not under duress; and, as we will see, this does not entail that a particular mental event is a component of voluntary action.⁴ On the other hand, according to a volitional account a particular kind of mental event is a component of voluntary action.

In neuropsychology, the volitional way of thinking about voluntary action is quickly becoming orthodoxy. As a testament to this, most neuropsychological research on voluntary agency is aimed at either examining (a) some sort of mental event thought to be an essential feature of voluntary action,⁵ and/or (b) whether certain bodily motions have such an event in their causal history as a way of examining whether we are actually voluntary agents, or just deceived if we think so.⁶ However, unless the volitional way of thinking is correct and the Aristotelian way wrong, most neuropsychological research on voluntary agency is misguided. And this illustrates the importance of examining our thinking about voluntary conduct—and what distinguishes forms of conduct, like Cain's killing of Abel, which we all naively think of as voluntary—before making determinations about whether science shows we never actually voluntarily act. In particular, it shows the importance of examining whether a particular mental event is a component of voluntary action.

So examining what distinguishes voluntary conduct is an essential step in the scientific study of voluntary agency. It is, however, widely

neglected. What is equally neglected is the more fundamental examination of what it is to perform an *act*, voluntarily or otherwise, in the way we typically think we do. Little has changed in the 50 years since Georg Henrik von Wright noted that, in discussions of this nature, ‘it is all too often taken for granted that it is clear what action [including voluntary action] *is*’.⁷ Most neuropsychologists, and many theorists in general, accept a volitional account of voluntary *action* without considering whether there might be a better option. A key aim of the present work is to criticise this oversight, and argue for an Aristotelian account. What we will see is that, if we actually examine our everyday thought about our conduct, including our *action*, the pattern our conduct takes, and how we come by the concept of voluntary conduct, what distinguishes that which, within the course of our everyday affairs, we think of as voluntary conduct aligns with an Aristotelian account. But—putting aside any inherent value in being correct—why does it matter whether a volitional or Aristotelian account is correct? Well, it may matter as much as it matters whether we are voluntary agents.

If a volitional account of voluntary action is correct—if voluntary movement is, in general terms, bodily motion that has a certain mental event in its causal history—then, as many in the field argue, there are neuropsychological grounds for thinking that voluntary agency is an illusion; and that willing, or choosing, is epiphenomenal. These grounds chiefly come from a line of research initiated by Benjamin Libet and his colleagues—what I will call *Libet-style experiments*—which indicate that brain activity leading to muscle contractions when one is, at least, thought to be acting precedes the mental event regarded as the ‘will’, or ‘choice’, to act.⁸ As Patrick Haggard claims, this research provides evidence that what we think of as choosing, or willing, to act amounts to no more than an experience that is ‘an immediate *consequence* of brain processes which prepare action’.⁹ Similarly, as Michael Gazzaniga claims, this line of research is ‘unlocking how and when the brain seem[s] to be making a person’s decisions for action’.¹⁰ The implication is, when we, at least, think a person is voluntarily acting, neural causal chains that produce the relevant movements seem to be initiated *before* the occurrence of the mental event volitional theorists think to be in the causal history of voluntary movement. Thus, the question in neuropsychology is predominantly no longer whether we have voluntary agency but rather, as Atsushi Sato states, ‘[w]hat kind of mechanism underlies’ the ‘feeling’ we do, i.e., ‘the sense of agency’?¹¹

An examination as to whether an Aristotelian or volitional account of voluntary action is correct is, therefore, relevant to addressing whether

4 *Human Agency and Neural Causes*

neuropsychology uncovers (a) brain activity required, and involved, when one exercises voluntary agency, or (b) brain activity that gives rise to the illusion we exercise voluntary agency. And if neural findings are taken to reveal that, in actuality, we do not wilfully act, or choose any of our conduct, this will have profound effects on the way we think about ourselves, and others. As a result, it will, also, have an effect on the way we live.

Psychological studies show that a lack of confidence in our ability to personally control whether certain things happen¹²—and, similarly, that telling people that voluntary agency is an illusion¹³—tends to have a negative impact on various aspects of life, including self-esteem, performance on cognitive tasks and the prevalence of prosocial conduct. Considering these findings, it seems virtually impossible to anticipate all the ways widespread acceptance of the view that voluntary agency is an illusion may negatively influence how, at least, *some* of us live.

To illustrate, if a person comes to believe that what she wills, and chooses, makes no difference regarding what will happen she may begin to see herself, as well as others, as more akin to an object that is manipulated by various forces. As a result, she may become less motivated, have lower self-worth, have lower expectations for herself, and be less concerned about how her conduct affects others.¹⁴ All of these dispositional shifts would have adverse influences on how she behaves towards others, and may even constrict what she is actually capable of doing.¹⁵ These shifts may make her less reflective about how her actions negatively impact others, and, thus, may make her more likely to engage in behaviour that harms others psychologically or physically. Based on the psychological studies mentioned in the above paragraph, these kinds of adverse effects would occur in a percentage of cases as a result of widespread acceptance of the view that voluntary agency is an illusion, which would have a negative impact on society.

Thus, while Sam Harris' recent prediction in *Free Will* that '[i]f the scientific community were to declare free will an illusion, it would precipitate a culture war far more belligerent than the one that has been waged on the subject of evolution'¹⁶ may be an exaggeration, his conclusion, that giving up the idea we have free will will be a net societal positive, is anecdotal¹⁷ and goes against psychological research that indicates otherwise. So—to answer the question about its relevance—it seems that the investigation as to whether neuropsychology, or neuroscience, shows voluntary agency, or free will,¹⁸ to be an illusion has practical, as well as theoretical, implications.

In this work I argue for an Aristotelian, rather than volitional, account of voluntary agency. I, further, maintain that our neural observations concerning voluntary agency can be plausibly interpreted in a way that is consistent with the idea we are voluntary agents in the Aristotelian sense I develop in this book. Finally, I situate the idea that we are voluntary agents within a broader, metaphysical framework. That is, I examine what is true about the world if we are, in fact, voluntary agents in the Aristotelian sense I develop here—a topic which should be of interest to those interested in assessing whether what we continue to find out about the world, ourselves included, is consistent with the idea that we are voluntary agents.

As we will see, not only will this work be a tale of two ideas about voluntary agency, it will also be a tale of two ways of thinking about the world. To apply current terms of art, it will be a tale of a certain ‘*reductive*’ way of thinking and a certain ‘*emergentist*’ way of thinking. As a first approximation, according to the first way of thinking, all that we experience, and all that takes place, can exhaustively be explained in terms of *subpersonal* causes. By contrast, according to the second, some of what takes place can only be explained in terms of what *people* do, and what *they* cause. As it turns out, the analysis I offer here indicates that the first, reductive way of thinking, which currently informs the way many theorists interpret neural findings, is inconsistent with the idea we are voluntary agents. However, this analysis also reveals reasons for being sceptical of this reductive way of thinking. For one thing, my analysis of our thought about what we do and *accomplish* (including what we accomplish through science and philosophy) provides indication that arriving at this reductive way of thinking as a scientific (and/or philosophical) *accomplishment* would be self-refuting. If this turns out to be true, it would indicate that this reductive way of thinking is unstable, and problematic; and that a shift is needed in the way neural observations are interpreted. This observation, along with the observation that this reductive interpretation hinges on speculative metaphysical commitments, will lead me to explore an emergentist way of interpreting our neural observations. I will argue that, though controversial, an emergentist interpretation is a plausible way of interpreting our neural observations that accommodates the idea we are voluntary agents.

Before beginning the project at hand, in Section 1.2, I shall provide more detail regarding the objectives and structure of this book. Then, in Section 1.3, I shall examine what have become traditional (compatibilist/incompatibilist) approaches to the topic of free will, as a way of positioning, and clarifying, my work here.

1.2 An outline of the work ahead

As indicated above, in this book I argue for an Aristotelian account of voluntary agency, and I make the case that our neural observations are consistent with the idea we are voluntary agents in this sense. To accomplish these objectives, I begin in *Chapter 2* by laying out the problem Libet-style experimental findings are said to pose for the idea that we are voluntary agents. I, then, discuss current arguments for why these findings do not pose a problem for the idea we are voluntary agents. As we will see, within the current debate concerning whether Libet-style experiments provide evidence against the idea we are voluntary agents, there is an underlying volitional assumption that has largely gone unexamined; roughly—the assumption is that voluntary movements have a certain mental event in their causal history. This assumption has to do with the very nature of voluntary conduct—what distinguishes various forms of conduct readily thought of as voluntary within the course of everyday life—and whether Libet-style experiments are the right way to go about testing whether we are voluntary agents.

After seeing that the argument that Libet-style findings bring the idea we are voluntary agents into doubt is predicated on a volitional assumption, I provide a critical examination of this assumption. This examination results in both conceptual and empirical grounds for being sceptical about volitional accounts. And by calling these accounts into question, I raise the possibility that, rather than providing evidence against the idea we are voluntary agents, Libet-style findings merely provide further evidence against a certain account of voluntary agency.

In *Chapter 3*, I question the motivation for upholding volitional accounts, and make the case that questions surrounding volitional accounts point toward the need for a focused analysis of our everyday thought, and talk, about our action in order to accurately assess what Libet-style findings, or any neuroscientific findings, reveal about voluntary agency. Von Wright is right: when drawing conclusions about such things, it is too often mistakenly taken for granted what an action is. To come to a correct account of voluntary agency, we must examine the patterns our conduct takes, our thought, and talk, about what we do and accomplish, and what distinguishes conduct we readily think of as voluntary. And we must come to a correct account of voluntary agency before we can accurately assess whether neuropsychology, or neuroscience, reveals that we are, in fact, not this kind of agent.

After making the case for an analysis of our thought, and talk, about our conduct, beginning in *Chapter 4*, I engage in this analysis. In brief,

my analysis reveals that if we actually perform actions in the way we typically think we do, then, when we do so, *we*, as *agents*, change a situation such that something other than what was going to obtain on a particular occasion obtains. My analysis, additionally, reveals that much of the conduct we think we engage in (e.g., purposive conduct, refraining)¹⁹ is conduct engaged in knowingly, or aware that one does so, when aware of being able to behave in alternative ways on that occasion—that is, *it is chosen*. However, I point out that, consistent with observations made in ‘embodied cognitive’ research, being aware of, or that, *x* should not be thought of as corresponding with a mental event of the kind that features in volitional accounts.

My analysis eventually leads to the conclusion that voluntary conduct is a subclass of chosen conduct. I develop this position largely on conceptual grounds; and by examining the way we typically think, and talk, about our conduct and what we accomplish. Though in places I explore unexamined territory, I draw from analytic work in the philosophy of action that reaffirms a broadly-conceived Aristotelian account of voluntary agency.

The analysis I offer in this work—though resulting in an account at odds with assumptions and views that shape current neuropsychological studies of voluntary agency—is not in competition with analysis performed by neuropsychologists. Rather, it is an attempt to fill a void created by a failure to bring an in-depth analysis of the way we typically think, and talk, about our conduct, in general, into the debate. And, as we will see in *Chapter 5*, such an analysis has, also, largely been neglected within the compatibilist/incompatibilist debate about free will—I will have more to say about this in Section 1.3.²⁰

I wrap up my analysis of voluntary conduct in *Chapter 6*. Then, in *Chapter 7*, I critically examine what I argue to be the most substantive *neuroscientific* threat to the idea we are voluntary agents, once properly conceived; namely—the prospect that our neural observations justify the currently predominant theory that *all* the relevant bodily changes, including neural changes, that occur whenever—so it is said—a person ‘acts’ can be given a sufficient causal explanation purely in terms of subpersonal causes (i.e., causes specified in the causal laws of the natural sciences). I call such theories *exhaustively mechanistic theories*; and, as my analysis reveals, such theories are inconsistent with the idea we are voluntary agents. But, as we will also see, exhaustively mechanistic theories are perpetuated by presuppositions brought to, and/or unsupported inferences drawn from, other neural observations.

To clarify what I do in Chapter 7: I do not claim that exhaustively mechanistic theories (or the models devised from them) are inconsistent with our neural observations. Rather, I argue that our neural observations, on their own, are equally open to being interpreted in a way that is consistent with our ordinary, or natural, way of thinking about our conduct, and what we accomplish by it; and in a way that is consistent with the idea that we are voluntary agents. In this case, though such theories may be warranted by other observations, exhaustively mechanistic theories are unwarranted by our neural observations, in and of themselves.

In addition to the above, in Chapter 7 I, also, provide an indication that once we step back and analyse our thought about what we accomplish through our scientific (or philosophical) endeavours we find that coming to an exhaustively mechanistic theory as a scientific (and/or philosophical) accomplishment would be self-refuting. Specifically, I provide novel reasons for thinking that putting oneself to developing, or otherwise coming to, an accurate theory, and accomplishing one's aim *by* coming to an exhaustively mechanistic theory, is, *itself*, the kind of phenomenon exhaustively mechanistic theories rule out. Hence, not only are exhaustively mechanistic theories unwarranted by our neural observations, we have a basis for rejecting these kinds of theories that supersede any basis we may think we have for upholding them. And this only becomes evident once we analyse our thought about what we do, and accomplish, including when accomplishing scientific feats. That is, once we perform the analytic work of Chapter 4.

If the arguments of Chapter 7 are correct, then not only are our neural observations consistent with the idea we are voluntary agents, but also upholding the kind of neuroscientific theories that are prominent in the field, and are inconsistent with the idea we are voluntary agents, is an unstable and implausible position. Thus, the observations made in Chapter 7 reveal serious problems that need to be addressed by proponents of exhaustively mechanistic theories.

Finally, in *Chapter 8* I bring this work to a close by approaching this question: If we are voluntary agents, what kind of world can we, and can't we, inhabit? I approach the question as to what kind of metaphysical framework is compatible with the idea that we are voluntary agents. This examination will help elucidate features of a plausible neural account of voluntary agency.

As I will argue in Chapter 8, while the currently popular position that all causes can, in principle, be specified in the causal laws of the natural sciences (e.g., physics, chemistry, neurobiology) is at odds with the idea that we are voluntary agents (once properly conceived),

a kind of emergentism, which seems to be growing in popularity, is not. Put roughly, the reason is, if we are actually voluntary agents, we as individuals, or persons, exercise distinct abilities and thereby cause change for which there is no sufficient subpersonal cause. Having made this observation, I wrap up this book by making the case that, though upholding the idea that we are voluntary agents would require a substantial shift from the way we currently interpret our neural findings, it seems that our current interpretations are shaped by a commitment to a speculative—and even problematic—metaphysical position that lacks empirical warrant. This observation leads me to develop a neuroscientifically informed emergentist account of human agency.

1.3 A different approach to a familiar topic: Undercutting the contemporary compatibilist/incompatibilist debate

The present book holds implications for the compatibilist/incompatibilist debate about free will. At the same time, it takes a different form than most contemporary philosophical works on free will. Nevertheless, in order to get our bearings before we set out on what may seem, at least within the contemporary landscape, a different trail, it may be helpful to, first, survey what have become common approaches to, and motivations for, examining free will. Doing so will help differentiate my approach and motivation. As we will see, while relevant to the current compatibilist/incompatibilist debate about free will, the issues I address here are different from those that preoccupy most contemporary philosophical discourse on free will.

In most contemporary philosophical approaches, the question as to whether deterministic ideas about the world threaten the idea we have free will is in the foreground.²¹ And there have commonly been two overarching motivations for this:

1. On the one hand, many theorists are motivated to approach the topic by the thought that our ideas about our own agency (and/or moral responsibility) should not, or need not, rest on whether, in the end, some form of determinism is shown to be true.²² (The most commonly considered form of determinism is what I will call *strict causal determinism*: that facts about the past and laws of nature necessitate all subsequent occurrences thereby ruling out all other logically possible occurrences; see Section 5.1.) These theorists argue for some idea of free will that is compatible with deterministic ideas about the world. Such views are called *compatibilist*.²³

2. On the other hand, some theorists are unconvinced by compatibilist ideas, and, thus, uphold an *incompatibilist* idea of free will; that is, an idea incompatible with deterministic ideas about the world. Such theorists can be further divided into two subgroups: (a) those who argue that we have free will in an incompatibilist sense, a position called *libertarianism*;²⁴ and (b) those who uphold an incompatibilist idea of free will and argue that there is good reason to think we do not have free will in this sense, a position called *hard determinism*.²⁵

Most philosophers who are currently engaged in the compatibilist/incompatibilist debate wisely note that the truth of determinism is hardly a foregone conclusion.²⁶ Rather, they take up the question of whether the idea we have free will is compatible with determinism as a metaphysical question concerning abstract concepts. Motivated by this metaphysical question, most theorists approach the topic of free will by immediately setting out to formulate some sort of *explicitly* compatibilist, and/or incompatibilist, idea(s) of free will to examine various arguments for, and/or against, these *abstract* ideas. And they do so without first undertaking an in-depth examination of our everyday thought, and talk, about our conduct and accomplishments.

In the present book I take a different approach; and I have a different focus. Here, I am particularly interested in developing a correct account of voluntary conduct; and, by and large, compatibilist/incompatibilist considerations are an afterthought. Compatibilist, or incompatibilist, concerns do not motivate the present work, the approach I take in it, or, ultimately, the account of voluntary conduct I develop. I do not set out to formulate either an *explicitly* compatibilist or incompatibilist account of voluntary agency, or of free will for that matter. Whereas it has become commonplace to stay in the stratospheric level of the abstract, I delve into what may, at first, seem the tedious minutiae of our everyday conduct and the way we think about it, and work my way up. I take this approach because I think, otherwise, crucial details concerning what voluntary conduct is are likely to be missed. (I, likewise, think that crucial details concerning what we typically think we *freely* and *wilfully* do—or do when we exercise free will—are likely to be missed.) Thus, in attempting to come to a sound conception of voluntary conduct, I, first and foremost, analyse the way we typically think and talk about our conduct, importantly including our action. I, then, let the chips fall where they may regarding the question of whether the idea that we actually engage in the kinds of conduct we typically think

we do is compatible with deterministic ideas. In this way, what contribution the present work may make to the compatibilist/incompatibilist debate concerning free will will come from a rather untraditional angle. It will come out of addressing preliminary, but, often, neglected, issues concerning ideas about our agency—that is, about human agency—in general (see Chapters 4 and 5).

Before expounding any further on my approach and motivation, I should point out that there is another recently popular way of approaching the topic of free will. In this approach, theorists and researchers examine what common ‘folk’, or non-specialists, think about free will, and whether a ‘folk psychological’ view of it is compatibilist or incompatibilist.²⁷ In the present book I, also, avoid this approach. I am not concerned with how most people think of free will, or, for that matter, with the idea of free will people come up with as a result of ‘introspection’; that is, as a result of paying attention to, and describing, the ‘experience’ of an instance of wilfully doing something. Let me explain why, as part of a more thorough explanation of how and why, this book will be different from the typical treatments of free will outlined above.

A substantial portion of this book is devoted to analysing our *thought* about what we are doing and accomplishing when we, at least, think we are doing such things as purposively acting and refraining. As we will see, this analysis reveals that actually doing and accomplishing much of what we think we do in everyday life has certain logical entailments (see Chapter 4). And these logical entailments may hold regardless of whether anyone—including any common ‘folk’, or theorists for that matter—recognises it as a result of introspection, or any other considerations. That something is true does not entail that anyone recognises its truth.

Thus, in short, my motivation from the outset is to develop a correct account of what we typically think we do and accomplish, including when we at least think we exercise voluntary agency. I am not motivated by compatibilist/incompatibilist concerns regarding free will. That being said, as we shall see, getting clear about what we typically think we do, *in general*, is crucial for accurately assessing whether our everyday thought about what we *freely* and *wilfully* do is compatible with *any* form of determinism; or any abstract ideas about the world for that matter. Further—and this is especially relevant to my central aim in this book—getting clear about what we typically think we do and accomplish, in general, will be crucial for coming to a correct account of voluntary conduct. And having a correct account of voluntary conduct

12 *Human Agency and Neural Causes*

is a crucial preliminary for accurately assessing whether neuropsychological, or neuroscientific, research provides evidence that we are not, in fact, voluntary agents. This is straightforward:

- (i) A neuropsychological, or neuroscientific, treatment of voluntary agency requires focusing on what is, at least, thought to be its exercise;
- (ii) We cannot focus on what is thought to be the exercise of voluntary agency without having an idea about what voluntary action is;

and

- (iii) If we start with an incorrect idea about what voluntary action is, we may focus on the wrong thing and draw errant conclusions;

(i)–(iii) are true even if, in the end, neuropsychological, or neuroscientific, research reveals we are not voluntary agents.

So I am, first and foremost, interested in coming to a correct account of what we typically think we do and accomplish. And my analysis leads me to an Aristotelian, rather than a volitional, position. Only, then, do I explore the implications for:

- (a) the popular metaphysical concern as to whether the idea we have free will is compatible with determinism;

and

- (b) take back up the burning question in today's intellectual climate as to whether neuroscience research reveals we are not voluntary agents.

As someone who has a background in neuroscience research, I think addressing what neuroscience research reveals pertaining to human agency has become particularly important and relevant.

With regard to (a), I argue that—as it turns out—the idea we behave in many of the ways we think we do (freely and wilfully or otherwise) is incompatible with strict causal determinism (see Chapter 5). At the same time, as will eventually come to light in the concluding chapter—Chapter 8—the idea we behave in these ways (freely and wilfully or otherwise) is compatible with the idea that nothing occurs by chance. The upshot is that, when considering whether we have free will, it does not matter whether we live in an indeterministic world (or, that is, in a

world containing chance occurrences). And this has implications for the compatibility/incompatibility debate. It is important to reiterate, however, that compatibilist/incompatibilist considerations do not motivate my analysis. Rather, they provide a wider metaphysical framework for understanding the outcome of my analysis.

With regard to (b), while certain classic neuropsychological problems for the idea we are voluntary agents disappear if we give up volitional accounts in favour of an Aristotelian account, this does not simply make all neuroscientific problems go away. Rather, accepting the Aristotelian account of voluntary conduct I develop here refocuses us on a different problem; namely—the problem posed by exhaustively mechanistic theories (as defined in Section 1.2). Thus, I devote a chapter to examining these theories—Chapter 7. As mentioned in Section 1.2, I examine reasons for thinking that our neural findings are equally open to being interpreted in a way that is consistent with the idea that we are voluntary agents. Here, it may be helpful to further clarify my point.

While I do not question whether we may be able to develop an exhaustively mechanistic theory that is consistent with our neural findings, I do, however, argue that exhaustively mechanistic theories are unwarranted by our neural findings, in and of themselves. I argue that a more ordinary, or natural, interpretation is available. As a first approximation, this more natural interpretation involves thinking of neural activities as part of the physical makeup of a *person* who is performing certain actions or activities—e.g., purposively acting, refraining—and *who* is, thereby, causing *certain* changes for which there is no sufficient underlying subpersonal cause. And, unlike exhaustively mechanistic interpretations, this interpretation is consistent with the ordinary, or natural, idea we have of ourselves as going through life purposively acting and refraining—an idea that, at least initially and naively, we bring to neuroscience, or any science. In this case, our neural findings do not, of themselves, warrant upholding an exhaustively mechanistic theory. But in addition to being unwarranted by our neural findings, as I mentioned in Section 1.2, it seems a real possibility that stepping back and examining a wider range of factors—factors including what we set out to *accomplish in and through scientific (and philosophical) endeavours*—reveals that coming to an exhaustively mechanistic theory as a scientific (and/or philosophical) *accomplishment* would be self-refuting. If this is in fact the case, something is wrong with exhaustively mechanistic ways of thinking. And—for reasons developed in Chapters 7 and 8—I think this is a plausible scenario that, at the very least, needs to be addressed by exhaustively mechanistic theorists.

An illustration may help further elucidate the argument outlined above. When staring at the classic ‘My Wife and My Mother-in-Law’ illusion (see Figure 1.1) we can either see the drawing as a picture of a young girl or of an old woman, but we cannot see the picture as both simultaneously. Further, seeing it either as a picture of a young girl or an old woman is not a result of failing to pay attention to the details. Rather, the details permit the image to be seen in both ways. However, if the image were presented in some wider context—for example, if there were figures in the background whose size and orientation made it clear that it was a picture of a young girl—then it would be clear what the drawing is a drawing of. In a similar way, our neural findings, *in isolation*, may be *equally* consistent with an exhaustively mechanistic way of thinking or, alternatively, with the way we ordinarily, and naturally, think about what we do and accomplish. However, when we step back and consider the wider context in which these findings are made—the



Figure 1.1 ‘My Wife and My Mother-in-Law’ Illusion

Source: German Postcard (1888), public domain.

context provided by the various phenomena of human life—it may very well become clear which way of thinking is accurate. It may become clear that when we are imaging the neural activities taking place within a conscious, fully functioning human being in action, we are observing parts of an agent who is purposively acting (or, on other instances, refraining or deliberating, etc.), and who is, thereby, causing *certain* changes for which there is no sufficient underlying subpersonal cause.

In summary, what I offer here is a work in philosophical psychology which brings philosophy of action to bear on the scientific examination of voluntary agency in a way that (i) shows that certain bold ‘scientific’ claims rely on, at least, questionable accounts of voluntary agency, and (ii) supports an Aristotelian account of voluntary agency. In the process, serious reasons for being sceptical of certain views of the world that are currently influential within neuroscience, and are in competition with the idea that we are voluntary agents, are uncovered. Additionally, observations are made that have implications for the compatibilist/incompatibilist debate about free will.

Index

A

acting, 181, 183, 184
 action, 3, 103, 104, 105–6, 108, 110,
 112, 114, 116, 117, 122, 124–5.
 See also voluntary action
 bodily motion. *See* bodily motion
 concept of, 64
 thinking about, 61–72
 action potential, 217n45
 agent, 7, 17, 116–17
 psychophysical, 132, 133, 134,
 141–2, 148, 154, 168, 170, 171,
 172, 175, 177, 179, 180, 182, 190
 voluntary, 5–7, 12, 13, 15, 16–17, 20,
 23, 25, 26, 29–30, 45, 50, 52, 55,
 59, 102, 104, 111, 116–17, 120,
 125, 128, 129, 131, 165, 167, 171,
 172, 174, 178, 180–2, 185–7, 189
 agent-causal account, 2, 17, 28–9, 31,
 62, 116, 167, 168, 170, 182, 186,
 189–90
 Alexander, Samuel, 170
 Alien (Anarchic) Hand Syndrome,
 33–4, 41
 alternative states-of-affairs, 54, 61,
 62, 67–72, 85, 91, 95, 96–9, 101,
 138–9, 142, 160, 161, 162, 164
 Alvarez, Maria, 92
 analytic approach to human agency,
 53–5
 Anscombe, Elizabeth, 53, 225n39
 anterior cingulate cortex, 145, 177
 antipsychotic drugs, 150
 Aristotelian account, 1, 3, 53, 57–85
 Aristotle, 57, 82, 108–10
 Assadi, Seyed, 145
 attention, 81
 awareness, 50, 72–85, 177
 attention vs., 81
 neuroplasticity, 81
 of one's conduct, 74–83
 pre-reflective states of, 81
 visual, 81

B

Bank, Gordon, 33
 Bargh, John, 212n37–213n44
 basic action, 31–7, 48
 Beckermann, Ansgar, 167
 behaviour, 58, 106–7, 108, 111
 Bennett, Karen, 220n85, 221n90
 Bennett, Max, 36
 bodily motion, 31, 32–3, 65–6, 105–6,
 116, 140–1, 180–1
 agent-causal account of, 2, 17,
 63–5, 105–6
 event-causal account of, 2, 17, 27–8
 intentional, 27
 readiness potential (RP), 25
 subpersonal causes of, 23
 unconscious processes related to, 24
 voluntary, 116
 Bok, Hilary, 99–100
 brain activities. *See* neural activities
 brain hemispheres, 19
 Broad, C. D., 169–70
 Brunswick, Egon, 199n70
 Burge, Tyler, 156–7

C

Cartwright, Nancy, 172–3, 184
 causal closure principle, 171, 173–4
 causal connections, 86–7, 208n3
 causal determinism. *See* strict causal
 determinism
 causal explanations
 for biological phenomena, 179
 for psychological phenomena, 179
 of behaviour, 158
 causal forces, 117
 Causal Inheritance Principle, 223n24
 causal laws, 87, 174–5, 208n5, 212n35
 causal powers, 143, 156, 184
 natural, 57, 171, 172–5, 180, 181,
 185, 187, 188
 non-aggregative, 171, 175, 178,
 180, 182, 190, 225n39, 226n47

- causal relations, 86–7, 185, 223n24, 226n49
- causative verb, 203n17
- cause, 5, 7, 9, 15, 22–3, 105, 116, 126–8, 134–44, 145, 154, 157–60, 171–2, 174–80
- change, 13, 17, 21, 29, 63–71, 76, 77, 81, 105, 110, 113, 116–17, 119, 127–8, 133–49, 154–5, 157–61, 171–91
- bodily, 7, 17, 29, 116–17, 127–8, 133–44, 148, 154, 157–8
- neural, 7, 117, 126, 127, 133–44, 146, 148, 154, 158, 172, 176, 177, 180, 185, 186
- choices, 59–61, 103–25. *See also* power to choose
- chosen conduct, 7, 55, 102, 104, 106, 107, 109–11, 115, 119, 125, 128, 181. *See also* voluntary conduct
- Clark, Sam, 22
- Claxton, Guy, 22
- Clayton, Philip, 179
- cognitive dysfunctions, 151
- compatibilist, 9, 88–101, 180, 182–3, 185–6, 188. *See also* free will
- complex neural systems. *See* neural systems
- compulsion, 106
- conduct, 58
- awareness of, 48–50, 188
- chosen, 110–11
- voluntary. *See* voluntary conduct
- conscious control, 24
- conscious decision, 17, 196n10
- conscious mental event, 17, 22, 24–5, 27, 29, 30, 31, 35, 39, 40–4, 46–7, 49–52, 59, 62, 72, 83, 84, 195n4
- consciousness, 137, 140, 151–3, 179
- conscious vetoes, 24–5, 43
- conscious willing, 19
- Consequence Argument, 88–9
- constitutive parts, 168, 170–1, 175, 176–7
- D**
- dACC. *See* dorsal anterior cingulate cortex (dACC)
- Davidson, Donald, 167, 168, 221n5
- daydreaming, 39
- decision, 201n5. *See also* deliberation
- decision-making, 48–9, 75, 80, 86, 90, 91, 95, 101
- Deecke, Lüder, 19
- Deliberation, 51, 60–1, 69, 82, 88, 98–101, 143, 201n5
- Dennett, Daniel, 99, 101, 162–3
- desire, 28, 47, 51, 59–60, 66, 90, 97, 98, 139, 158
- Di Nucci, Ezio, 212n32
- dispositional shifts, 4
- DLPF. *See* dorsal lateral prefrontal cortex (DLPF)
- dopamine receptors, 146, 151, 153
- dorsal anterior cingulate cortex (dACC), 177
- dorsal lateral prefrontal cortex (DLPF), 126, 176
- downward causation, 171, 191, 223n24
- drugs, 150–1
- Dupré, John, 158, 175
- duress, 106–11
- E**
- ecological validity, 30, 199n70
- Eimer, Martin, 19–21
- electroencephalography (EEG), 18
- embodied cognitive research, 7, 81, 118, 120, 189
- emergence, 5, 166–74
- and downward causation, 171
- of mind, 166, 167, 169–70
- of voluntary agency, 177–9
- strong, 180–1
- weak, 226n47
- entity, 17, 64, 66, 80, 115, 116, 119, 131–4, 141–2, 144, 148, 154, 167, 168–71, 172, 175–8, 181, 182, 183, 191
- epistemic distinction, 49
- epistemic openness, 99, 204n39
- event-causal account, 2, 17, 27–9, 43, 53, 61–2, 64–6, 71, 116, 204n32
- exhaustively mechanistic theories, 13–14, 128–31
- contradictory/self-refuting, 155–65
- subpersonal causes, 5, 7, 9, 16, 22–3, 126–8, 134–45, 154, 157–60, 171–2, 174–80

F

- Flanagan, Owen, 24
 Frankfurt, Harry, 90–1, 111–15
 Frankfurt-type examples, 111–15
 free will, 21, 61, 86–101
 and consciousness, 24
 as an illusion, 4
 compatibilist/incompatibilist debate
 about, 9–15, 86, 88–101, 180,
 182–3, 185–6, 188
 compatibilist objections, 96–8
 Hierarchical Mesh Theory of, 90–1
 incompatibilist implications,
 88–95
 Reasons-Responsive Theories of,
 93–4
 Fried, Itzhak, 21
 functional Magnetic Resonance
 Imaging (*fMRI*), 21
 functional muscle synergies, 135

G

- Gazzaniga, Michael, 3, 126–7, 191–2
 Gleason, Curtis, 16
 Goldman, Alvin, 159
 Gomes, Gilberto, 196n10
 Grabenhorst, Fabian, 149

H

- Hacker, Peter, 36, 183
 Haggard, Patrick, 3, 19, 126
 Hampshire, Stuart, 103
 hard determinism, 10
 Hare, R. M., 221n3
 Harris, Sam, 4, 22
 Harré, Rom, 115
 Hart, H. L. A., 103
 Hierarchical Mesh Theories, 90–1
 Hofer, Carl, 207n1
 Honderich, Ted, 25
 human agency, 11, 61–72, 134–44
 analytic approach to, 53–5
 Hyman, John, 63–4

I

- illusion, 4, 115–25
 incompatibilist, 9–15, 88, 95–6
 indeterministic world, 178, 186,
 195n25

- intention, 26–8, 29, 65–6
 intracellular neurobiochemical
 activities, 146–51, 153
 intransitive verb, 63, 203n17
 ion flux, 217n45, 218n53

J

- Johnson, Mark, 184

K

- Kenny, Anthony, 57, 58, 115, 193n4,
 199n71, 202n13, 202n14, 202n7,
 210n6, 210n8, 211n17, 211n21,
 214n17
 Kim, Jaegwon, 159, 167–8, 172
 Kirsch, Irving, 22
 Koch, Christof, 23
 Kornhuber, Hans, 19

L

- Lakoff, George, 184
 language, 58, 201n12
 language-use, 30, 53
 lateralized readiness potential (LRP),
 19–20
 laws
 causal, 87, 174–5, 208n5,
 212n35
 trans-ordinal, 169
 layered model of universe, 168
 libertarianism, 10
 Libet, Benjamin, 3, 16. *See also*
 Libet-style experiments
 Libet-style experiments, 3, 16–42,
 116–17
 Locke, John, 115
 Lowe, E. J., 25–6
 LRP. *See* lateralized readiness potential
 (LRP)
 Lucas, J. R., 108
 Lynn, Steven, 22

M

- Madden, Edward, 115
 Malcolm, Norman, 157–9
 McKenna, Michael, 87
 medial prefrontal cortex (mPFC), 145,
 176
 Mele, Alfred, 24, 26–7

- mental event of willing, 18–21,
31–43, 52
 ecological validity of, 30–1
mental state, 54, 72–84, 118, 179–80
metabotropic receptors, 145–6
midbrain, 137, 151, 153
Mill, John Stuart, 169
mind-body supervenience.
 See supervenience
moral responsibility, 91, 93–6
 and Frankfurt-type examples, 111–15
Morgan, C. Lloyd, 221n3
mPFC. *See* medial prefrontal cortex
 (mPFC)
muscle contractions, 3, 18–22, 116,
127, 140, 143, 181
 functional muscle synergies, 135
mutative verb, 203n17
My Wife and My Mother-in-Law
 illusion, 14–15
- N**
neural activities, 13, 18, 19, 20–1,
80–1, 116–17, 124, 134–54, 174,
176–80, 184–5, 189, 190, 218n55,
226n48
neural firing patterns, 149–52
neural mechanisms, 120, 126–65
 and subpersonal causes, 5, 7, 9,
16, 22–3, 126–8, 134–44, 154,
157–60, 171–2, 174–80
neural systems, 144–55
neurophysiological systems, 152
neuroplasticity, 81, 136
neurotransmitters, 151
New Dispositionalism, 96
non-aggregative causal power, 171,
175, 178, 180, 182, 190, 225n39,
226n47
- O**
Oakley, David, 24
O'Connor, Timothy, 187, 222n6
OFC. *See* orbitofrontal cortex (OFC)
orbitofrontal cortex (OFC), 176
- P**
Paley, William, 166
Papineau, David, 223n23
- Pearl, Dennis, 16
performing
 a mental act, 44, 49, 59, 62, 66, 72,
77, 80–1, 84, 96, 105, 108, 116
 an act, 26, 31, 35, 48, 49, 51–2, 67,
69, 71, 72, 79, 83, 133, 140, 141,
142, 143, 144, 171, 172
phlogiston, 43, 200n87
physical constitution, 116, 133
Platt, Michael, 126
Popper, Karl, 166–7, 169, 192
posterior parietal cortex (PPC), 176–7
powers
 causal. *See* causal powers
 rational, 57–8
 two-way, 57–8, 61
power to choose, 54, 55, 58, 59, 61,
62, 72–85, 102, 103, 104, 105,
115–17, 119, 122, 124, 128, 129,
132, 142, 155, 161, 164, 178, 180,
182, 184–8
 and downward causation, 171
 and voluntary agency, 106–11
PPC. *See* posterior parietal cortex
 (PPC)
pre-reflective states of awareness, 81
pre-supplementary motor cortex
 (pre-SMA), 19
primary motor cortex (M1), 19, 20,
139
priming studies, 117–24
proximal decision, 26–7, 199n57
psychological attributes, 131–2
psychological phenomena, 156–7
psychophysical agent, 132, 133, 134,
141–2, 148, 154, 168, 170, 171,
172, 175, 177, 179, 180, 182, 190
psychophysical causation, 171
purposive action, 33, 54, 62, 69, 76,
77–8, 93, 98, 160
- R**
rational powers, 57–8
readiness potential (RP), 18–19
Reasons-Responsive Theories, 93–4
reductive, 5, 143, 180, 188
refraining, 13, 54, 57, 62, 69–71,
74–7, 83, 108, 203n12
Rolls, Edmund, 149

Ruben, David-Hillel, 202n8
 Rushworth, Matthew, 126, 145

S

Sato, Atsushi, 3
 semicompatibilism, 194n23
 sense of agency, 3
 sensual desires, 59–60
 Sirigu, Angela, 41
 SMA. *See* supplementary motor cortex (SMA)
 social psychology, 117–24
 speculative amendments, 37
 Sperry, Roger, 170
 Steward, Helen, 27, 28, 43, 49
 stochastic neural activity, 135, 153, 184, 215n27–8, 225n38
 Stout, Rowland, 60
 Strawson, P. F., 35, 94
 strict causal determinism, 9, 86–8, 207n1
 strong emergence, 180–1
 subpersonal automatic processes, 16
 subpersonal causes, 5, 7, 9, 16, 22–3, 126–8, 134–45, 154, 157–60, 171–2, 174–80
 sufficient causal explanation, 7, 127–8, 154, 158, 177, 181, 187
 supervenience, 167–8, 221n3, 222n5, 222n7
 supplementary motor cortex (SMA), 19, 21

T

Tancredi, Laurence, 21, 126–7
 top-down causation. *See* downward causation
 transitive verb, 63, 203n17
 trans-ordinal laws, 169
 traumatic brain injury, 151
 two-way power, 57–8, 61

U

unconscious processes, and bodily motion, 24
 unilateral neglect syndrome (UNS), 81
 universe

depictions of the, 180–92
 layered model of the, 168
 UNS. *See* unilateral neglect syndrome (UNS)
 urges, 18–19, 26–7, 36, 40, 47

V

van Inwagen, Peter, 86, 88–9
 Velleman, J. David, 27–8
 ventromedial prefrontal cortex (VMPF), 149
 VMPF. *See* ventromedial prefrontal cortex (VMPF)
 volitional accounts, 2
 agent-causal, 17
 conceptual and empirical criticisms of, 30–42, 46–52
 event-causal, 17
 voluntary action, 3–4, 38, 103, 104, 105–6, 108, 110, 112, 114, 116, 117, 122, 124–5
 Aristotelian account of, 1–2
 bodily motion. *See* bodily motion
 intentional, 26
 volitional account of, 2–3, 16–17, 25, 26, 29, 30–1, 35, 37, 38, 39, 41–2
 voluntary agency, 2
 emergentist account of, 174–80
 power to choose, 106–11
 voluntary conduct, 11–13, 39–42, 53–5, 103–11, 199n70
 active/passive distinction, 40
 Aristotelian account of, 1–2, 106–11
 misconceptions about, 46–52
 voluntary prisoner example, 114–15
 von Wright, Georg Henrik, 3, 53–4, 63, 203n12

W

Wallace, R. Jay, 94–5
 wants, 46–51, 59–61
 weak emergence, 226n47
 Wegner, Daniel, 127
 Wittgenstein, Ludwig, 35, 53, 103, 108, 200n77, 201n8, 210n2, 210n11
 Wolf, Susan, 91–3
 Wright, Elwood, 16