



OXFORD

Consciousness &
Moral Responsibility

NEIL LEVY

Consciousness and Moral Responsibility

Consciousness and Moral Responsibility

Neil Levy

OXFORD
UNIVERSITY PRESS

OXFORD

UNIVERSITY PRESS

Great Clarendon Street, Oxford, OX2 6DP,
United Kingdom

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide. Oxford is a registered trade mark of
Oxford University Press in the UK and in certain other countries

© Neil Levy 2014

The moral rights of the author have been asserted

First Edition published in 2014

Impression: 1

All rights reserved. No part of this publication may be reproduced, stored in
a retrieval system, or transmitted, in any form or by any means, without the
prior permission in writing of Oxford University Press, or as expressly permitted
by law, by licence or under terms agreed with the appropriate reprographics
rights organization. Enquiries concerning reproduction outside the scope of the
above should be sent to the Rights Department, Oxford University Press, at the
address above

You must not circulate this work in any other form
and you must impose this same condition on any acquirer

Published in the United States of America by Oxford University Press
198 Madison Avenue, New York, NY 10016, United States of America

British Library Cataloguing in Publication Data

Data available

Library of Congress Control Number: 2013954185

ISBN 978-0-19-870463-8

Printed and bound in Great Britain by
CPI Group (UK) Ltd, Croydon, CRO 4YY

Links to third party websites are provided by Oxford in good faith and
for information only. Oxford disclaims any responsibility for the materials
contained in any third party website referenced in this work.

Preface

On a cold February evening in 1870, a porter at a Kentucky hotel attempted to awaken a man named Fain, who had fallen into a deep sleep in the lobby. At first the porter thought Fain might be dead, so hard was it to rouse him, but eventually Fain stirred. Witnesses report hearing him ask the porter to leave him alone, but the porter wanted to close the lobby and dragged Fain to his feet. Fain drew a gun and shot the porter three times. He was later convicted of manslaughter, but subsequently acquitted on appeal (*Fain v. Commonwealth*). His defense argued that Fain was sleepwalking and not conscious of what he was doing.

As the Fain case illustrates, sometimes agents perform morally serious actions while apparently not conscious of what they are doing. In some of these cases, the agent appears to lack consciousness altogether. These cases are dramatic, puzzling, and relatively rare. Far more common are cases of agents who are normally conscious, but fail to be conscious of some fact or reason which nevertheless plays a role in shaping their behavior. A candidate for a job might be rejected, for instance, because he nonconsciously reminds a search committee member of her ex-husband. In other cases, which might be more common still, agents are conscious of facts that shape their behavior, but conscious neither of how, nor even *that*, those facts shape their behavior. There is a vast amount of experimental evidence for this kind of effect. Consider, for instance, a recent experiment by Schnall et al. (2008). They seated participants at either a clean desk or a dirty desk while the subjects assessed the permissibility of various actions. For those participants who scored in the upper half of a scale measuring consciousness of one's own body, being seated at a dirty desk led to stronger moral judgments. Participants were aware of the dirtiness of the desk and of their disgusted response to it, but not of how their disgust influenced their moral judgments.

Situations in which our behavior is shaped by facts of which we are either not conscious, or not conscious of their effects on our behavior, are ubiquitous. Whether other agents are deliberately manipulating our behavior (in supermarkets, for instance) or by happenstance, our behavior is constantly and continually shaped and modulated, if not caused, by nonconscious stimuli. Sometimes, these stimuli shape our behavior in morally serious ways. Again, these cases are puzzling.

Until recently, theories of moral responsibility paid little heed to these cases. Implicitly, they often seemed to assume some kind of Cartesian mental transparency, whereby agents are able to access all the contents of their minds. In the past decade, however, a number of philosophers have turned their attention to these kinds of cases as they develop and test theories of moral responsibility. Given that the evidence from cognitive science clearly demonstrates that minds are not transparent and that consciousness of all the reasons to which we respond is obviously too demanding as a condition of moral responsibility, many philosophers now advocate much more modest conditions. Several prominent theorists argue that we ought to reject the implicit assumption that consciousness—at least consciousness *of* the nature of our actions or of the reasons to which we respond—is a necessary condition of moral responsibility. This book argues that these philosophers are wrong. Consciousness of key features of our actions is a necessary condition of moral responsibility for them. In particular, moral responsibility requires that agents be conscious of the *moral significance* of their actions. This fact, I will argue, excuses agents from moral responsibility in a variety of cases: not only in the bizarre cases involving nonconscious agents, but also in far more common cases in which agents fail to be conscious of key facts shaping their behavior. The argument I shall present is very significantly empirical. It rests on claims about consciousness and its functional role, and especially the claim that consciousness is required for the integration of information. Because the view rests in very significant part on empirical claims, it is open to empirical refutation. I believe that the account of consciousness offered here is well supported by the available data. Further research

may show it to be false; it is worth remarking, however, that in that case it will most likely be replaced by a view according to which behavior responds to information without much in the way of broad integration of its contents. Such a view would be so at odds with the folk psychological roots of concepts like moral responsibility that we would have little idea whether agents should be held responsible or not. I believe that the view presented here offers us the best chance of vindicating those aspects of folk psychology most directly tied to our concepts of agency and responsibility. Further, I think that the view is most likely true.

The book has six short chapters. In Chapter 1, I sketch some of the background motivating the denial of what I shall call the consciousness thesis: the thesis that consciousness of at least some of the facts that give our actions or omissions their moral significance is a necessary condition of moral responsibility. I briefly review both scientific evidence for the ubiquity and power of nonconscious processes, and philosophical arguments against the consciousness thesis. In Chapter 2, I discuss the challenge to the consciousness thesis provoked by the work of Benjamin Libet and Daniel Wegner. I show that this challenge is simply irrelevant to moral responsibility: it makes no difference whether or not consciousness has the powers they contend it lacks. I then set out the consciousness thesis in a little more detail. I delineate both what kind of consciousness is at issue, and what the content of the relevant conscious states must be. In Chapter 3, I set out the global workspace account of consciousness, a theory of consciousness that has received a great deal of scientific support, and defend it against objections. In Chapter 4, I advance an account of the role that consciousness plays in behavior, building on the global workspace theory. I argue that because consciousness plays the role of integrating representations, behavior driven by nonconscious representations is inflexible and stereotyped; only when a representation is conscious can it interact with the full range of the agent's personal-level propositional attitudes. This fact, I argue, entails that consciousness of the facts to which we respond is required for these facts to be assessed by and expressive of the agent herself, rather than merely a subset of her attitudes.

Chapters 5 and 6 examine how two popular and plausible theories of moral responsibility fare in the light of the consciousness thesis. Chapter 5 focuses on the view that moral responsibility requires expression of the agent's real self, or their evaluative agency, as I prefer to say. I argue that only when attitudes that imbue actions with moral significance or which otherwise reflect that significance are *globally broadcast* are those actions expressive of the agent's evaluative agency, because only under these conditions is the moral significance assessed for consistency and conflict with the personal-level attitudes constitutive of the real self. I argue that this rules out moral responsibility both for actions performed by nonconscious agents and for actions caused, in more normal conditions, by agents who happen not to be conscious of attitudes that impart moral significance to an action. I also argue that nonconscious attitudes lack the kinds of contents that could plausibly underwrite attributions of moral responsibility for actions that fall short of expressing evaluative agency.

In Chapter 6, I turn to control-based accounts of moral responsibility. I focus on Fischer and Ravizza's (1998) notion of guidance control, since it is an undemanding form of control: if agents who are not conscious of the moral significance of their actions do not exercise guidance control over them, they would seem unlikely to satisfy any plausible control condition for moral responsibility. I argue that nonconscious agents do not exhibit the regular reasons-receptivity required for guidance control, and that conscious agents who happen not to be conscious of the moral significance of their actions thereby fail to exercise guidance control over the morally relevant features of their actions. This chapter concludes with a few brief remarks on George Sher's (2009) account of moral responsibility, which is explicitly tailored for rejecting the consciousness thesis. Sher's account has fewer empirical commitments than the expression and control views, so rebutting it requires more than citing the data. Instead I tackle the account philosophically, presenting three reasons why I find it unsatisfactory.

In a brief conclusion, I address the arguments of philosophers who accept the consciousness thesis, but who maintain that the

pervasiveness of nonconscious processes entirely rules out moral responsibility. I demonstrate that agents in the kinds of cases that these philosophers have in mind satisfy the consciousness thesis, and that there is a significant difference between these cases and those in which agents are excused due to the absence of consciousness. Finally, I offer a few remarks on methodology, suggesting that progress on the issues canvassed here requires a deep engagement with the sciences of the mind. We cannot rely on intuitions alone, because intuitions may be generated by processes that fail to track the truth. I hope not only to convince readers of the claims made here, but also to provide an illustration of one way (not the only way) in which philosophy of action should be done and thereby encourage others to follow.

The view I defend here occupies a part of logical space that is otherwise sparsely inhabited. Most people who think that consciousness is necessary for moral responsibility, or for agency at its best, are motivated by a high regard for consciousness. They may hold, for instance, that agents are to be identified with their conscious states. Those philosophers who deny that consciousness is necessary for moral responsibility are motivated, on the contrary, by a conviction that the contents of consciousness tell only a very small part of the story of who we are. I concur wholeheartedly with the latter claim, and am at the very least tempted to deny the former: consciousness is never more than a tiny sliver of our mental life, and the contents that happen to become conscious may not be especially significant for who we are. Consciousness is necessary for direct moral responsibility, I claim, not because of what it *is*, but because of what it *does*. The contents that constitute our identity are broadly distributed in the mind, and the vast majority of these contents are at any one time nonconscious. Consciousness is a tiny, and very frequently unrepresentative, portion of our mental life, but consciousness enables the distributed mechanisms that constitute agents to play a coordinated and coherent role in agency.

Let me end this preface by addressing a question that might arise for anyone who has read my previous work. In *Hard Luck* (Levy, 2011(a)), I argued that (for reasons independent of consciousness) no one is ever morally responsible for anything. I have not changed my mind about

that claim. Given that fact, it might reasonably be asked why I am bothering with inquiring whether consciousness is a necessary condition of moral responsibility. Atheists don't tend to worry themselves over the attributes which God would have, were there to be a God,¹ so why would I worry about the necessary conditions for the application of a concept that I believe can never justifiably be applied at all? To that question I have four responses. First, exploring the commitments of our concepts is worth doing in its own right (as a matter of fact, I know atheists who work professionally on the attributes of God and who do not seem embarrassed by this fact). Second, given that I doubt that my arguments for the claim that no one is ever morally responsible for anything will convince many people, but I do think that establishing that consciousness is a necessary condition for moral responsibility is an easier task and one that might meet with greater approval, the task is worthwhile. It will lead to fewer people being unjustly held morally responsible (I don't know whether it is a greater injustice to be held morally responsible when one fails to satisfy the consciousness thesis or when one does not, but given that it *is* an injustice, it is worth the fight: whatever our other views, we all agree that—other things being equal—the fewer people unjustly held morally responsible the better).

Third, even though I do not believe that anyone is ever morally responsible, I do believe that people and their behavior can be morally assessed: we can distinguish between better and worse actions, and vicious and virtuous individuals. In defending the consciousness thesis, I aim to further our ability to make such assessments. Only when our actions are expressions of our selves can we be appropriately identified with them, such that we can be assessed on their basis, I will argue. This is of practical importance to us in our everyday lives. We often want to know whether the actions and omissions of our friends,

¹ Of course, atheism may be justified by claims about the nature of God; an atheist might argue that God is an impossible being because His properties are mutually incompatible, for instance. But I don't think that the conditions of moral responsibility I explore here are mutually inconsistent or otherwise impossible to actualize, so I can't avail myself of this kind of justification of my current project.

work mates, intimate partners, and so on, reveal something morally deep about them and their concerns; whether (for instance) their forgetting that it is our birthday, or their blurting out of a confidence, reveals something important about their values and commitments. I shall claim that only when agents satisfy the consciousness thesis do their actions and omissions express their attitudes; for this reason, the consciousness thesis matters for our moral lives even in the absence of a sufficient condition for moral responsibility. Fourth, and finally, though I believe that blame and the practices that build upon blame are not justified, it remains legitimate and necessary to impose certain kinds of sanctions on agents in certain circumstances, for reasons of deterrence and incapacitation. But in order to know which agents ought to be sanctioned, when, and how, we need to know whether their behavior was reasons-responsive; consciousness of the reasons for which we act, I shall argue, is needed for reasons-responsiveness. Further, we need to know whether their behavior expressed their attitudes, and which attitudes it expressed, because this knowledge will allow us to predict their future behavior. When agents satisfy the consciousness thesis, I shall argue, their actions express their genuine attitudes and they possess the capacity to exercise personal-level control over their behavior; hence this condition tracks other factors which remain important, even in the absence of moral responsibility.

I have not attempted in these pages to offer a sufficient condition of moral responsibility (obviously, since I do not believe that there is a sufficient condition of moral responsibility), nor even to adjudicate between rival conceptions as to which is most plausible. Nor have I sought to do opponents of the consciousness thesis justice: I put aside important differences between their views and do not engage in any real depth with their often subtle arguments. I have attempted, instead, to lay out an argument for the consciousness thesis as pithily and as persuasively as possible, and to reply to the major objections to the view. Readers who seek deeper engagement with the views of opponents of the consciousness thesis, as well as additional and independent arguments against them, may find both in some of my other work (Levy, 2011b; 2012; 2013;

forthcoming). I hope the tighter focus that the development of a positive thesis allows somewhat makes up for the occasional crudities entailed.

Even if no one is ever morally responsible, the consciousness thesis matters. It is important whether it is true. I aim to demonstrate that it is true.

Acknowledgments

Many people gave me extremely helpful comments on portions of the manuscript of this book; some were generous enough to read and comment on the whole thing. The resulting book has been improved immeasurably, though I am painfully aware that I have not been able to respond adequately to all the comments. I am especially grateful to Tim Bayne, Simon Rippon, Markus Schlosser, Nick Shea, Richard Tamburro, and Alan White. Portions of the argument of the book at various stages of its development were presented to audiences at the Australasian Association of Philosophy conference, Wollongong; the Zentrum für interdisziplinäre Forschung, Bielefeld; the Florey Institute of Neuroscience and Mental Health, Melbourne; the University of Leiden; the University of Oxford; The University of Latvia; the University of Valencia; and Southern Methodist University, Dallas. I am grateful to the audiences at these talks for very helpful feedback. Three external referees for Oxford University Press provided uncommonly helpful and perceptive comments on the manuscript, showing that constructive criticism is not yet a dead art. I also owe a debt of gratitude to Peter Momtchiloff for his patience with the project and for selecting such high-quality referees, and to Jack Howells for expert copyediting.

Though the central argument defended here is new, in defending and developing it I have drawn from material previously published. Chapters 2, 3, 5, and 6 contain paragraphs drawn, with more or less considerable modifications, from the following articles: “Expressing Who We Are: Moral Responsibility and Awareness of Our Reasons for Action,” *Analytic Philosophy* 52 (2011), 243–61; “Neuroethics,” *Wiley Interdisciplinary Reviews: Cognitive Science*, 3 (2012), 143–51; “The Importance of Awareness,” *Australasian Journal of Philosophy*, 91 (2013), 211–29; and “Consciousness, Implicit Attitudes and Moral Responsibility,” *Noûs*, forthcoming.

Contents

1. Does Consciousness Matter?	1
2. The Consciousness Thesis	14
3. The Global Workspace	38
4. What Does Consciousness Do?	70
5. Consciousness and the Real Self	87
6. Consciousness and Control	109
<i>Concluding Thoughts</i>	131
<i>References</i>	136
<i>Index</i>	151

1

Does Consciousness Matter?

This book will be an extended argument for what I hereby dub *the consciousness thesis*. The consciousness thesis is the thesis that consciousness of some of the facts that give our actions their moral significance is a necessary condition for moral responsibility. I will not attempt to offer a *sufficient* condition of moral responsibility. Rather, my approach will be to show that only when we are conscious of the facts that give our actions their moral significance are those actions expressive of our identities as practical agents and do we possess the kind of control that is plausibly required for moral responsibility. These conditions—the expression of our identities as practical agents and the possession of a kind of control—are the two primary contenders for sufficient conditions of moral responsibility available today. Each has been powerfully and plausibly defended by a number of philosophers. I do not seek to add to that literature; rather, I seek to convince the participants in these debates that they should accept the consciousness thesis, no matter which theory they profess.

I take the approach of showing that the two major competing accounts of moral responsibility are both committed to the consciousness thesis for two main reasons. First, showing that both popular theories are committed to the thesis has obvious dialectical advantages. It allows me largely to avoid the contentious debate between these theories; this is obviously desirable, since it makes it more likely that my conclusions will be acceptable to both sides. There is, I think, less at stake in the competition between these views than their defenders think: once each is modified to take account of the considerations advanced here, the views will not conflict over who is responsible