

PHILOSOPHY



Traditional and Experimental Readings

Fritz Allhoff
Ron Mallon
Shaun Nichols

New York • Oxford
OXFORD UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further Oxford University's objective of excellence in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Copyright © 2013 by Oxford University Press, Inc.

For titles covered by Section 112 of the US Higher Education Opportunity Act, please visit www.oup.com/us/he for the latest information about pricing and alternate formats.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016
www.oup.com

Oxford is a registered trademark of Oxford University Press

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, electronic, mechanical, photocopying, recording, or otherwise, without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data

<to come>

Printing number: 9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

4.3 “Brain Damage, Mind Damage, and Dualism”



MARK PHELAN, ERIC MANDELBAUM, AND SHAUN NICHOLS

MARK PHELAN (1976–) is an Assistant Professor of Philosophy at Lawrence University of Wisconsin. Previously, he was a Postdoctoral Researcher in the Philosophy and Cognitive Science Departments at Yale University. His work focuses on interactions between linguistic pragmatic phenomena and cognitive processing. His papers focus on such topics as figurative language, intentional action, group minds, the nature of evidence, and experimental research on pragmatic processing. Eric Mandelbaum (1981–) is a Mind/Brain/Behavior Postdoctoral Fellow at Harvard University. His contribution to this essay was supported by an Oxford Martin Research Fellowship at the University of Oxford. His work focuses on cognitive architecture, belief acquisition, belief storage, implicit racism, and unconscious inference. Shaun Nichols (1964–) is a Professor of Philosophy at the University of Arizona and a leading experimental philosopher. In this selection the authors argue that brain damage poses a fatal objection to Cartesian dualism over and above the interaction problem for dualism.

Living is a rather good thing. Living longer is (generally) better. Asked if they would rather keep living after today, most people would agree that they would. However, there might come a point when living doesn't seem like such a wonderful thing. If you find yourself at age 110 bedridden, incapable of controlling basic bodily functions, and in constant agony, living may not look so good anymore. From this we can suppose that, if what came with immortality were the agonizing ravages of our physical bodies, we might not want immortality. If living in agony isn't good, then living in agony forever would be infinitely worse.

But perhaps the most important thing about a person—the mind—is protected from the inevitable decay of the biological organism. According to a family of influential philosophical views—a family we'll characterize with the label “dualism”—the mind is an immaterial soul that can continue existing after biological death.¹ In the European philosophical tradition, dualism

is most prominently associated with Rene Descartes, who argued that the mind is an immaterial substance that communicates with the physical brain. As a result, he maintained that “the decay of the body does not imply the destruction of the mind.” And this, he thought, “afford[s] to men the hope of a future life.” For, “while the body can very easily perish,” Descartes writes, “the mind is immortal by its nature” (Descartes 1641/1984, 10). Although this venerable philosophical view of the mind may seem abstract and detached from everyday life, it resonates with what most people think, according to the psychologist Paul Bloom. Bloom writes: “most people . . . believe that the soul can survive the complete destruction of the body” (p. XXX, this volume). By contrast, other prominent philosophical theories of the mind make it difficult to see how this is possible. Physicalist accounts of the mind, accounts that equate mental states to neurological states of the brain, identify the mind with a decaying organism, and

M. Phelan, E. Mandelbaum, and S. Nichols, “Brain Damage, Mind Damage, and Dualism.”

thus seem to preclude the continued existence of a mind after (biological) death.

While dualism appears to provide for the possibility of the immortality of the mind,² it faces the problem of *interaction*—if the mind is an immaterial substance, how can it causally interact with physical bodies (Elisabeth, p. XXX, this volume)? Some philosophers are optimistic that dualism can address the interaction problem (see, e.g., Lycan 2010). Although it is intuitively puzzling how an immaterial soul can push around physical objects, one reply is to acknowledge that there are lots of puzzling phenomena in the world, especially when it comes to what causes what. For present purposes, we want to set aside the interaction objection itself. The effects of brain damage, we will argue, pose an acute explanatory challenge over and above the interaction problem.

Let's explore the problem of brain damage by considering the famous case of Clive Wearing (CW). CW lived a normal life for almost 50 years, when he contracted a form of viral encephalitis. The virus not only caused "anterograde amnesia," a disability where one can no longer form new memories, but it also caused a fairly severe form of "retrograde amnesia," a disability where one cannot recall events and facts from before the onset of the neurological trauma. Brain scans on CW show that the areas associated with memory (e.g., the hippocampus) were largely destroyed. Having both deficits, CW's case covers two clear types of amnesia: an inability to recall events and facts prior to a traumatic incident and an inability to form new memories after a traumatic incident.³

According to dualism, the mind and the brain are distinct substances, with the mind existing wholly outside the physical world. Yet the effects of brain damage are quite difficult to explain if the mind is on an entirely separate plane from the body, for brain damage seems to have drastic effects on mental states like memory. So, if the mind is a nonphysical soul, not susceptible to the "corruption of the body," then how could brain damage—a physical process—damage the mind?

Following the scientists who work on patients like CW, we have said that CW's case involves the

dramatic loss of specific memories of the past as well as critical mental capacities for forming memories. Moreover, we have followed the scientists in suggesting that the damage to the mind was caused by brain damage. But the dualist might maintain that this is all a mistake. While it might appear that CW has mind damage, the dualist may insist that this is not really the case, instead alleging that his mind remains completely undamaged. For the dualist to defend this view, though, she needs to provide some explanation for the *appearance* of mental damage. Dualism's most promising strategy for explaining the apparent mental damage rests on the idea of a brain-mind interface.⁴ Suppose the brain is (in part) a kind of modem or information router, one that sends physical inputs from the body to the immaterial mind and receives transmissions from the mind, transforming them into electrochemical responses that the body can use.⁵ If the brain and the mind normally communicate with one another, then some failure to send information in one direction or the other may explain the behavior of a brain-damaged patient like CW. This strategy, if defensible, would preserve the possibility that the mind is not damaged by brain damage, in keeping with dualism.

Let's clarify the specifics of this strategy by focusing on CW's retrograde amnesia. CW, like most amnesiacs, has selective amnesia. For example, he can remember that he has children, but he cannot remember their childhood; he can remember how to play piano, but he cannot remember any past episodes of playing the piano; he can remember that he is married, but he can no longer remember getting married. How can the dualist explain these phenomena? If there has been no mental damage, then why can't CW retrieve his memories?

To preserve the integrity of the nonphysical soul, the dualist might say that the problem is restricted to the communication between the (undamaged) mind and the damaged brain. However, it isn't enough for the dualist simply to say, "It's a communication problem." Since we are, for the moment at least, granting that interaction between the immaterial mind and the physical

body is possible, we are willing to grant that such communication is possible. But we are now comparing the merits of the physicalist explanation of these deficits to the merits of a dualist account. Since the physicalist supposes that the mind is, at the very least, affected by the brain, the physicalist can point to connections between the brain damage and the corresponding mental deficits. Such an explanation can give a fairly detailed account of how such damage occurs, and such an account would be explanatorily superior to the dualist merely saying, “It’s a communication problem.” So, for the dualist, more needs to be said about the nature of the problem. There are two obvious options available. One possibility is that while the mind is perfectly fine, the brain garbles input that comes to it from the mind (see the “dualist distorted input view” on Figure 1). The other possibility is that the brain produces the wrong output sent to the mind (see the “dualist distorted

output view” on Figure 1). Let’s consider each of these explanations in turn.

Let’s begin with the distorted input view by focusing on a specific example. When asked, “Do you remember your wedding?” CW claims that he does not. The dualist might maintain that CW has the memories of his wedding stored in his immaterial mind, but the neurological damage causes his response to come out wrong: instead of saying, “It was in a beautiful setting, with lots of friends there to celebrate with us...,” he just says, “I don’t remember.” But this dualist response leaves one wanting: why would the brain damage only selectively affect some of CW’s responses? Compare: if he is asked whether he is married, he will respond affirmatively. So, in order for the dualist response to be viable, it must explain the selective effects, in particular, why CW’s answer “comes out wrong” in some cases but not others; otherwise, such a response

Note: Please provide the Figure caption.

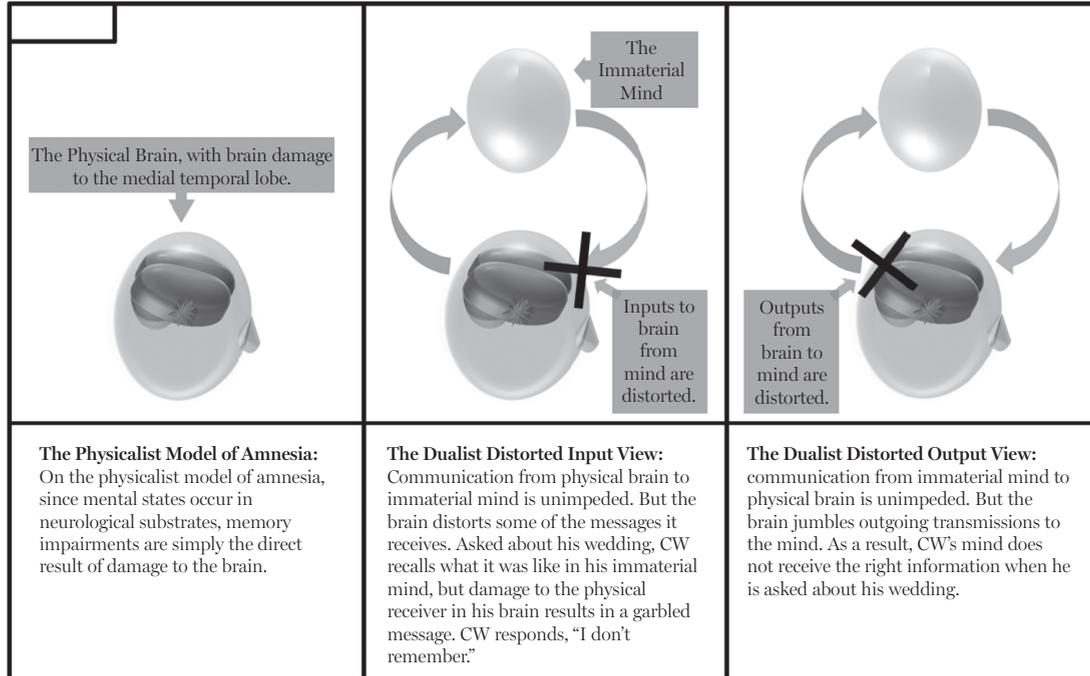


Figure 1

lacks sufficient explanatory detail to compete with the physicalist explanation. Perhaps, instead, the dualist will respond by saying that the mind does not send out the correct memory; instead, it just sends out instructions to say, “I don’t remember.” But again, why should the soul send out these instructions only selectively? And, anyway, why send out instructions to say *that*?

Perhaps it is better for the dualist to explain CW’s problems by looking at things from the other direction, by supposing these are problems with the outgoing line, not the incoming one (see the “dualist distorted output view” on Figure 1). On this take, the dualist would respond that because the line is down, the signals from the brain can’t get through to the mind to allow the formation of new memories; in other words, the anterograde amnesiac brain can’t send the signals out to the mind to get stored in memory. When someone asks CW, “Do you remember your wedding?” the sounds get transmitted into his ears, but when his brain-modem sends the signals to his mind, it just comes out as jumbled garbage. Hence, his mind does not receive the right information in order to form the question. However, this dualist response makes it puzzling why CW answers, “I don’t remember.” If you asked someone on the street how to get to the Eiffel Tower, a reasonable response would be, “I don’t know” (assuming she didn’t know). If, on the other hand, all she heard was mumbled garbage, a reasonable response would be, “Come again?” But CW never responds by saying, “I didn’t hear you” or “I don’t understand the question”; rather, he seems to perfectly understand the question and just cannot answer it, for he can’t ascertain the requisite knowledge required to answer it.

So, neither the distorted input view nor the distorted output view provides a very satisfying explanation of *the mere appearance* of mind damage resulting from a damaged brain.⁶ Perhaps the dualist has no option but to countenance *actual damage to the mind* stemming from physical damage to the brain. This raises further

explanatory burdens for dualism, to be sure. But, more important for the present context, it also seems to defeat dualism’s promise to maintain our minds in the face of biological decay. If the dualist explains CW’s responses as most scientists do, in terms of actual mind damage (i.e., the dramatic loss of specific memories of the past along with critical mental capacities for forming new memories), then the dualist admits that the death of the brain can deprive whatever soul lives on of past memories of its human life, as well as the capacity to store new experiences of the afterlife.⁷ This means the immortal soul will lose the love we feel when we recall when our children were young, the pride in our past accomplishments, the happiness we experience when we remember times spent with our dear friends. . . . If the dualist has to embrace mind damage, then she must accept that the joy of remembrances of such past events dies with our physical body. And one might ask what good eternal life is for a soul that loses that.

Notes

1. This is to paint with a very broad brush. There are many different types of dualism that one might hold. In this essay, we focus on what is called “substance dualism,” according to which the mind consists in an altogether different type of substance than the body, the mind being a mental substance and the body being a physical one. More specifically, we focus on a substance dualism that allows for causal interaction between the physical and mental substances. “Property dualism” is a much weaker position than substance dualism. The property dualist thinks that there is just one type of substance in this world, physical substance, but two ultimate types of properties: physical and mental properties. The mental properties are analyzed as basic and fundamental properties that somehow emerge from the brain. We focus on substance dualism because, unlike property dualism, substance dualism (at least in its traditional form) maintains that the death of the brain wouldn’t cause the death of the mind.
2. The issue of immortality is complicated by questions about what makes a person the same across time.

For instance, on some views, the persistence of the soul would not necessarily mean the persistence of the self. For this article, we set aside those additional complications (but see Chapter 7 on the self]).

3. For further discussion of CW see Sacks' (2007) *New Yorker* article.
4. Indeed, Descartes himself had an elaborate theory about how the mind and the brain communicated with one another (see Reading 4.1, this volume).
5. Of course, had one not granted a solution to the interaction problem, the problem would arise here, too, if one wondered how physical signals can affect a nonphysical mind (and how a nonphysical mind can send a physical signal).
6. The dualist might attempt other explanations. For instance, the dualist might reply that the brain isn't a single modem, but is instead a slew of modems communicating with the mind for different purposes, and this is what explains CW's pattern of behavior. But it's not clear that this response is adequate. Suppose we were to ask the dualist, how many modems does the brain have exactly? As we mentioned earlier, CW seems to be both unable to recall events and facts from prior to his illness and unable to form new memories after his illness, but there are people who have just one kind of deficit or the other. So, the dualist would have to posit two different modems to account for these two separable problems. And it seems that the dualist can't stop here. Numerous other kinds of brain damage would seem to require other modems: some traumas render people unable to identify tools; some cause people to fail to identify animals; some cause people to be unable to recognize who people are from seeing their face. . . . Should we assume there is a modem for each type of deficit? We started by assuming that maybe the brain is the modem for the mind and now we are asked to suppose that the brain contains multitudes of modems. Perhaps the most important challenge for the *multiple modem* line of response is to articulate some underlying principle that would justify positing a modem at any point. It would be *unprincipled* for the dualist just to posit another modem every time there is some datum his theory didn't predict. Alternatively, the dualist might avoid unprincipled multiplication of modems by pointing to neural complexity. Like a computer, the brain has complex, interrelated processes. Though your laptop has just

one Wi-Fi card installed, various patterns of local damage to different internal processes can render it unable to connect to the Internet in distinctive ways. Maybe the dualist would suggest something similar is going on in various cases of brain damage. But this response still faces the challenge of explaining the *specific pattern* of CW's deficits. Assume, as this response suggests, that localized damage renders CW's brain unable to transmit questions about past events to his immaterial mind. That still leaves it a mystery why CW responds, "I don't remember," to a question his mind never received. These two examples show how difficult it is for the dualist to articulate an account of the mind-brain relation that explains actual patterns of behavior that follow brain damage.

7. One might also think that our memories are what underwrite our personal identity (Parfit 1984). Thus, it may not even make sense to say that it would be *your* soul that survived death.

References

- Bloom, P. "The Duel Between Body and Soul." *New York Times*, September 10, 2004. Reprinted in this volume.
- Descartes, R. *Meditations on First Philosophy*, trans. John Cottingham, Robert Stoothoff, and Dugald Murdoch. In *The Philosophical Writings of Descartes*. Cambridge: Cambridge University Press, 1641/1984.
- Lycan, W. "Giving Dualism Its Due." *Australasian Journal of Philosophy* 87, no. 4 (2010): 551–563.
- Parfit, D. *Reasons and Person*. Oxford: Oxford University Press, 1984.
- Sacks, O. "The Abyss: Music and Amnesia." *New Yorker*, September 24, 2007. http://www.newyorker.com/reporting/2007/09/24/070924fa_fact_sacks?currentPage=all (accessed December 1, 2010).

Study Questions

1. Why is the brain damage objection a different objection from the interaction objection?
2. Why might the dualist distorted input view be seen as unappealing?
3. Why might the dualist distorted output view be seen as unappealing?