2.1 Seduced by Tradition

Daniel C. Dennett

Gazzaniga's essay provides a useful elementary overview of the ways in which physics has moved away from Laplace's vision of determinism and how the sciences in general have moved to a more nuanced appreciation of the relationships between multiple explanatory levels—subatomic, atomic, molecular, cellular, organismic, personal, social. He quotes an apt observation by Philip Anderson (1972): "The ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe." Neuroscientists, he says, have been slow to appreciate the idea of emergence and often persist in trying to couch their interpretations in the language of what I have called *greedy* reductionism (Dennett, 1995, pp. 81–82) as contrasted with good reductionism (which amounts, in Gazzaniga's terms, to *no magic*). I would add that an even greater foible of the neuroscientists is their uncritical reliance on the obsolete categories of prescientific traditions, what Wilfrid Sellars (1962) called the manifest image.

In fact the point that emerges most forcefully for me from Gazzaniga's essay is not what he intended to demonstrate, but nevertheless something he should be happy to have shown, because it is an important first step in clearing away the ancient presuppositions that make the free will issue so resistant to dissolution: The traditional ways of thinking about these phenomena—about decisions, about selves or minds, about conscious control—are relentlessly seductive. Even after we have glimpsed and appreciated better perspectives thanks to advances in science, we find ourselves being drawn back to the old habits of thought, trying to find, in our modern scientific picture of the brain, our dear old friends from long ago: the ego, the immortal soul, impervious to causal influence, wellspring of choices so free that even God could not predict them in advance. They aren't there to be found, of course, and Gazzaniga does a good job of sketching what is in our brains instead of these items, but much more important, they

don't *have* to be there for us to be agents with the sort of freedom that is a prerequisite for moral responsibility, for genuine authorship of our deeds and misdeeds. Many scientists are still succumbing to the temptation to *assume*—for it is never carefully argued for—that if these antiquated notions are illusory (if "free will is an illusion" as so many of them put it), then so is our moral agency. This theme in recent public pronouncements by leading scientists, especially neuroscientists, is deplorable, and Gazzaniga has attempted to expose the flaws in this thinking: "In what follows, while the goal will be to challenge the very concept of free will, the concept of personal responsibility remains untouched. The idea outlined below is that a mechanistic concept of how the mind works eliminates the need for the concept of free will."

Gazzaniga inadvertently shows, however, how nearly irresistible the categories are, by lapsing into them himself. Consider this, from his final summary: "The course of action taken appears to us as a matter of 'choice,' but the fact is, it is the result of a particular emergent mental state being selected by the complex interacting surrounding milieu."

That phrase, "but the fact is," suggests that the "particular emergent mental state being selected" is *not* a choice, especially not a *free* choice, and Gazzaniga underlines this suggestion with his closing line: "Our interpreter then claims we freely made a choice." The almost-invited inference is that our interpreter in the left hemisphere fools us, convinces us that we made a free choice when in fact we didn't. This needs to be challenged. There are indeed times when we fool ourselves—when our interpreter fools us if you like—into thinking we are making a free choice when in fact we are being manipulated by some other agent, or when the "complex interacting surrounding milieu" is seriously deranged by delusion or other cognitive pathology. When, on the other hand, we have our wits about us, and are not massively misinformed or otherwise manipulated, then *there is no important sense* in which the outcome of all the interactions in the many levels or layers of "machinery" is not a free choice. That's what a free choice *is*! It's the undistorted, unhindered outcome of a cognitive/conative/emotive process of exquisite subtlety, capable of canvassing the options with good judgment and then acting with fairly full knowledge of what is at stake and what is likely to transpire.

We say we try to choose our actions as the best options "all things considered," but of course we cannot consider *all* things, and not even all *relevant* things, but we can accomplish a fair approximation of that investigation, thanks to all the cognitive machinery and its interlocking systems of emotive control. The fact that our decision systems are "automatic,

deterministic, modularized, and driven not by one physical system at any one time but by hundreds, thousands, and perhaps millions" (Gazzaniga, this volume) does not show that they are not just what the doctor ordered for those of us who want to be responsible agents. Artificial intelligence has not yet produced an artificial decider as robust and reliable as your average well-informed citizen, but at least we're beginning to understand the sorts of subsystems and "moving parts" from which to construct such a thing. And nothing we have yet learned in cognitive neuroscience and related fields suggests that any and all such systems would have to be defective.

That, however, is the impression many neuroscientists are conveying to the general public. Tom Wolfe, an acute and sardonic taster of the winds of cultural change, puts it dramatically in his aptly titled essay, "Sorry, but your soul just died" (2001, p. 100): "The conclusion people out beyond the laboratory walls are drawing is: *The fix is in! We're all hardwired!* That, and: *Don't blame me! I'm wired wrong!*"

Wired wrong? What would it be to be wired right? *That* is a question for cognitive scientists to answer, and when they don't even address it, they are succumbing to the temptation to go along with—and even endorse—the mythology that anything that is "wired," that is mere "machinery," couldn't possibly have the kind of freedom required for moral responsibility.

Above I called this theme in the public statements of neuroscientists deplorable. Isn't that a little strong? Well consider this thought experiment: Once upon a time a brilliant neurosurgeon said this to a patient on whom she had just performed an implantation procedure in her shiny high-tech operating theater:

> The device I've implanted doesn't just control your obsessive–compulsive disorder; it controls your every decision, thanks to our master control system which maintains radio contact with your microchip 24 hours a day. In other words, I've disabled your conscious will; your sense of free will henceforth will be an illusion.

In fact she had done no such thing; this was simply a lie she decided to tell him to see what would happen. It worked; the poor fellow went out into the world convinced that he was not a responsible agent, but rather a mere puppet, and his behavior began to show it: He became irresponsible, aggressive, negligent, indulging his worst whims until he got caught and put on trial. Testifying in his own defense, he passionately protested his nonresponsibility, because of the implant in his brain, but nobody believed him, and the neuroscientist, when called to testify, admitted she'd said it,

"but I was just messing with his mind—a practical joke, that's all. I never thought he'd believe me!" It really doesn't matter whether the court believed him or her, sentenced him or her; either way she ruined his life with her ill-considered assertion, robbing him of his integrity and shutting down his conscience. In fact, her false "debriefing" of her patient actually accomplished nonsurgically much of what she claimed to accomplish surgically: She disabled him for life.

If we agree that she is responsible for this dire consequence, what shall we say about the neuroscientists currently filling the media with talk about how their science shows that free will is an illusion? Are they not risking doing the same, by mass production, to all the people who take them at their word?

Michael Gazzaniga has been keenly aware of this responsibility and has been a vigorous and resourceful leader in the effort to clarify these delicate issues for the general public. And yet even he can be lulled into honoring—at least not challenging—a presupposition that should be banished.

Notice that in my story the neuroscientist didn't *say* she had destroyed his moral responsibility; she "just" said she had rendered his free will illusory, and it was he who drew the further dire conclusion. That is, however, the natural conclusion for lay people to draw, and overcoming that presumption is an uphill battle. In addition to Gazzaniga, there are some philosophers—notably John Martin Fischer—who claim that free will is not necessary for moral responsibility, but it is a hard sell, given such familiar locutions as "You signed the contract of your own free will" and "Since you weren't coerced, but acted of your own free will, you are responsible …" Here the issue is not factual, or metaphysical, but tactical, a matter of good pedagogy or diplomacy. Two apparently wildly different claims turn out to *mean the same thing*, once you understand how the terms are being used:

A. Neuroscience shows that we don't have free will, but we can nevertheless be responsible for our actions under normal conditions.
B. Neuroscience shows that we do (under normal conditions) have free will, but free will turns out to be quite different from what prescientific folks thought it was.

Wilfrid Sellars (1962) characterizes the task of philosophy as negotiating the problematic traffic of thought between what he calls the manifest image (the everyday world of colors and sunsets and people and solid objects and free will) and the scientific image (of proteins, atoms, electrons, and the like). The free will issue is above all an opportunity to find paths

between these two perspectives that are not deeply and even catastrophically misleading. Prescientific understanding of free will is tethered to two conditions: It is supposed to be incompatible with determinism, and it is supposed to be a necessary condition for responsibility. Something has to give, but each tether is a powerful intuition that must either be honored in how we use the term or denied with careful argument. (The "libertarians"[1] try to honor both tethers, with the result that they end up defending mysterious doctrines of "agent causation" which defy any scientific understanding.)

Whenever philosophers or scientists have to negotiate the foggy chasm that separates the manifest image from the scientific image, tactical issues arise. Given what ordinary folks think color is, you *might* say that *color is an illusion*; nothing turns out to be colored in the nonscientific sense of the term. And nothing is solid, not really, not *through-and-through*. And dollars are an illusion, and so is software! Not a very convincing tactic. Or you could say colors are perfectly real, but quite surprisingly different from what you thought they were, and similarly for dollars and software—and free will. (I know from long and frustrating experience that no matter how many times I say that consciousness is real, but not what you think it is, many "clever" commentators assure us that the title of my book should have been *Consciousness Explained Away* or *Consciousness Denied*, that what I'm actually saying is that consciousness isn't real. Well in *their* sense of consciousness, they're right! Consciousness is not a magical extra ingredient in the physical world, and not a nonphysical ingredient either.)

Free will, similarly, is beset with "magical" connotations, a "tiny miracle" in the words of one libertarian philosopher (who did not mean them as disparagement!). If we decide, tactically, to let that tether fix our understanding of the term, then Gazzaniga is right "to challenge the very concept of free will, [while] the concept of personal responsibility remains untouched." But then he has to work harder on the positive account of responsibility, showing in detail how (shockingly, to the layperson) it does not require free will. And that task will be made much more difficult by the chorus of other neuroscientists who say (with tradition, and hence with scant argument if any) that responsibility does ("of course") require free will, and therefore we are *not* responsible, not really.

The tactical alternative, followed here by me, is to examine the requirements of responsibility and argue that nothing in neuroscience to date shows that those conditions are not met by normal adults, and that it is thus legitimate, honoring the other tether, to conclude that we have discovered what free will really is—nothing magical, nothing miraculous—

but just something rather unlike the inflated notion of tradition. As with our everyday concepts of color and solidity, we have to bend a bit, and take on a more informed and sophisticated concept to see that it is a perfectly real phenomenon. One of the chief advantages of this tactical alternative is that it allows us to honor the intuitively compelling contrasts between normal people who do have free will and people suffering from pathology, victims of manipulation, and so on, who do not. If we go the other way, we have to admit that a normal person and a psychotic or manipulated "puppet" person are really no different in the free will department. Neither has free will.

Since cognitive science can contribute a lot to our understanding of the differences between normal people and pathological decision makers *and why the differences are important for morality and law*, the second tactic is thus much to be preferred since it allows us to couch our *scientific image* explanations in terms of which conditions diminish or obliterate free will and which do not. That is what the law will need, and common understanding as well.

Note

1. Unfortunately, philosophical tradition has its own meaning of "libertarian," which has nothing to do with the more widespread political meaning(s). A free will libertarian is one who claims that free will is incompatible with determinism, is necessary for responsibility, and exists.

Moral Psychology

Volume 4: Free Will and Moral Responsibility

edited by Walter Sinnott-Armstrong

A Bradford Book
The MIT Press
Cambridge, Massachusetts
London, England

This book was set in Stone Sans and Stone Serif by Toppan Best-set Premedia Limited, Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data

Moral psychology / edited by Walter Sinnott-Armstrong.
v. cm.
"A Bradford Book."
Includes bibliographical references and index.
Contents: v. 1. The evolution of morality : adaptations and innateness—v. 2. The cognitive science of morality : intuition and diversity—v. 3. The neuroscience of morality : emotion, disease, and development.—v. 4. Free will and moral responsibility
ISBN 978-0-262-19561-4 (vol. 1 : hardcover : alk. paper)—ISBN 978-0-262-69354-7 (vol. 1 : pbk. : alk. paper)—ISBN 978-0-262-19569-0 (vol. 2 : hardcover : alk. paper)—ISBN 978-0-262-69357-8 (vol. 2 : pbk. : alk. paper)—ISBN 978-0-262-19564-5 (vol. 3 : hardcover : alk. paper)—ISBN 978-0-262-69355-4 (vol. 3 : pbk. : alk. paper)—ISBN 978-0-262-02668-0 (vol. 4 : hardcover : alk. paper)—ISBN 978-0-262-52547-3 (vol. 4 : pbk : alk. paper)
1. Ethics. 2. Psychology and philosophy. 3. Neurosciences. I. Sinnott-Armstrong, Walter, 1955–
BJ45.M66 2007
170—dc22

10 9 8 7 6 5 4 3 2 1