

Neuroethics for the New Millennium

Commentary

Adina Roskies

Department of Linguistics and Philosophy
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

The past several months have seen heightened interest in the intersection of ethics and neuroscience. In the popular press, the topic grabbed headlines in a May issue of *The Economist* (2002) and was featured in a New York Times editorial (Safire, 2002). Professional societies were a step ahead, staging several meetings devoted to ethics and neuroscience since the beginning of this year. In January 2002, *Neuron* and the AAAS sponsored a symposium entitled “Understanding the Neural Basis of Complex Behaviors: The Implications for Science and Society,” which brought together a panel with expertise in the neurosciences, policy, ethics, and the law to discuss the recent advances in the neurosciences and their potential implications for science and society. The Royal Institution in London sponsored “Neuroscience Future” in March (see conference summary by J. Waldbauer on NMTF conference website [URL below]), and in May the Dana Foundation, in collaboration with Stanford University and UCSF, sponsored a conference boldly entitled “Neuroethics: Mapping the Field” (NMTF). The name “neuroethics” implies such a field exists, an “unexplored continent lying between the two populated shores of ethics and of neuroscience,” in the words of Al Jonsen, an organizer of the NMTF conference.

My aim here is to delineate what I see as the basic structure of this nascent field and to lay out some of the fundamental questions with which it is concerned. The views here have been informed by the presentations and discussions of numerous people who participated in these conferences, but also strongly reflect my personal perspective on what the field of neuroethics should be.

It is evident that neuroethics will overlap substantially with traditional issues in biomedical ethics. For instance, much of the recent work in ethics spurred by the Human Genome Project will be applicable, with perhaps slight modification, to some neuroethical problems. But if there is to be justification for identifying and promoting neuroethics as a new and important field, it ought not be merely a subdivision of bioethics, with issues and answers parallel to those that arise in other areas of biomedical research. The intimate connection between our brains and our behaviors, as well as the peculiar relationship between our brains and our selves, generate distinctive questions that beg for the interplay between ethical and neuroscientific thinking. The motivation for the newfound interest in bringing together neuroscientists, ethicists, journalists, philosophers, and policy makers arises from the intuition that our ever-increasing understanding of the brain mechanisms underlying diverse behaviors has unique and potentially dramatic implications for our perspective on ethics and for social justice. These are the issues that warrant the

introduction of a new area of intellectual and social discourse.

As I see it, there are two main divisions of neuroethics: the ethics of neuroscience and the neuroscience of ethics. Each of these can be pursued independently to a large extent, but perhaps most intriguing is to contemplate how progress in each will affect the other.

The Ethics of Neuroscience

The ethics of neuroscience can be roughly subdivided into two groups of issues: (1) the ethical issues and considerations that should be raised in the course of designing and executing neuroscientific studies and (2) evaluation of the ethical and social impact that the results of those studies might have, or ought to have, on existing social, ethical, and legal structures. Let me call, for convenience, the first the “ethics of practice,” and the second the “ethical implications of neuroscience.” For the most part, the ethics of practice is where traditional bioethics, as applied to neuroscience, resides. It includes familiar issues like optimal clinical trial design, guidelines for use of fetal tissues or stem cells or cloning, privacy rights to results of testing for neurological disease, and so on. However, the ethics of practice includes some questions peculiar to neuroethics. For instance, in a liberal democratic society such as ours self-determination is highly prized, and hence the importance of informed consent is central to medical practice and medical ethics. But neurodegenerative diseases and psychiatric disorders may impair cognition so that informed consent, as generally conceived, may be impossible. What guidelines should be in place for treatment or experimental participation in these cases? We also take it for granted that when making medical decisions, patients will choose what is in their best interests. Some disorders of brain chemistry, such as depression, defy such an assumption. Who should wield executive power when the subject cannot be counted on to choose what is best for his or herself?

The second subdivision of the ethics of neuroscience, the “ethical implications of neuroscience” is the area of neuroethics that is truly novel, and perhaps the most ripe for advancement. Its aim is to investigate the implications of our mechanistic understanding of brain function for society, and it will require integrating neuroscientific knowledge with ethical and social thought. Advances in neuroscience have the potential to create, and to remedy, serious social inequities. How we use our knowledge will shape our society. How, as we learn more and more about how the brain controls behavior and the causes of mental dysfunction, are we to reconcile this new knowledge with the social structures that allow our society to run more or less smoothly? For instance, it has been suggested that a large proportion of inmates on death row may have damaged or injured brains. If careful epidemiologic studies establish that this is the case, how should our views about moral and legal responsibility change, if at all, to accommodate this surprising fact? In the future (but not currently!), it may be possible to use noninvasive imaging techniques to determine whether a person is lying. There is some

indication that such imaging technology could be used to distinguish real from false memories. Perhaps even further in the future we could, with some degree of certainty, diagnose behavioral dispositions, motivations, or beliefs. In what cases can such information be used ethically? What are the privacy issues associated with thought? What are the consequences of reliable, but not perfect, diagnostic techniques? If someone knows that he or she is at some risk for, for example, a psychotic episode, should he or she be held legally responsible for actions undertaken while delusional in virtue of not having prevented the episode? How should decision-making proceed in the face of probabilistic predictions of behavior? Will the results of some predictive studies become self-fulfilling prophecies? Once our technology provides us access to the full spectrum of physiological states underlying behavior, will our common practice of identifying certain behaviors as normal or abnormal have meaning? Will our knowledge prompt us to redefine what "normal" is? Such issues will undoubtedly arise as our technology and understanding of human cognition improves, and it would be well to have thought them through before we are faced with them in the flesh.

It is easy to see how consideration of the ethical implications of neuroscience will affect issues generally thought to be in the realm of the ethics of practice. Criteria for life and death are currently linked to gross generalizations about brain function. Better understanding of developmental and cognitive processes may allow us to refine these notions to better identify life and death, biological benchmarks important for a number of policy-related issues such as abortion, termination of support, etc. Other definitions are extremely relevant to the daily lives of many: what is considered a disability, what will insurance pay for? As we understand more fully the varieties of ways in which brains operate, it will be an increasing challenge to define these terms in a way that does not marginalize, but nonetheless protects, the disabled, the at-risk, and the disenfranchised. At the same time, we need to be aware of the risk of overextending these concepts and of mitigating diversity through medication. The pharmaceutical advances that will surely stem from neuroscientific ones will present the following question: when can drugs be ethically used to enhance normal capacities, rather than to just treat deficits? What effect will our policies have on existing social and economic disparities? How can regulations be enacted to promote fairness and equality? What might the consequences for our society or another be if we fail to safeguard fair and equitable access to such enhancements? We will also have to come to terms with the charge that chemical and technological enhancements make us less human or "post-human" (Fukuyama, 2002) and perhaps revisit the question of what it is to be human.

The Neuroscience of Ethics

The second major division I highlighted is the neuroscience of ethics. Traditional ethical theory has centered on philosophical notions such as free-will, self-control, personal identity, and intention. These notions can be investigated from the perspective of brain function. Although the neuroscience of ethics today is far less developed than the ethics of neuroscience, and may not progress as quickly at first, it will be the area with truly

profound implications for the way ethics, writ large, is approached in the 21st century.

Already there are signs of a surge in interest in investigating the brain bases of moral cognition, and such studies are bound to burgeon in the coming years. How are decisions made in the brain? How are values represented? How are ethical decisions similar to or different from other types of decisions? Many thinkers have assumed ethical reasoning to be a variety of rational thought. But recent evidence suggests that emotions play a central role in moral cognition (Damasio, 1995; Greene et al., 2001). Does this undermine the view of ethics as rational or instead undermine the long-cherished division between reason and emotion? How will a better understanding of the biological basis of moral cognition and behavior modify our philosophical ethical framework? How will it affect ingrained notions of rationality and its importance to human existence?

Many of us overtly or covertly believe in a kind of "neuroessentialism," that our brains define who we are, even more than do our genes. So in investigating the brain, we investigate the self. What is the neural representation of "self" dependent upon? Is personal identity a brain-based notion? What consequences for our concepts of personal identity will alterations of the self-defining parts of us have? Some current interventions (and undoubtedly more in the future) will be such as to perhaps improve the health and functioning of the patient, but perhaps at the expense of altering the brain chemically or mechanically. Will certain medical or technological therapies change who we are? What are the ethical implications of such changes? Will we have to weigh the costs of biological death against continued life but destruction of our selves? Advances in neuroscientific research in relevant areas may change the very fabric of our philosophical outlook on life.

Although neuroscience is unlikely to answer metaphysical questions about determinism, it can certainly alter our perceptions of them. As our predictive grasp of complex behavior improves, how will the bolstered sense of the brain as a deterministic machine affect or undermine our notions of free will or of moral responsibility? Is self-determination, a driving concept in today's bioethics, merely an unscientific fiction? If not, what is the biological basis for it, and if so, what notions should replace it? What brain structures are essential to self-control? How, and to what extent, can the role for self-control in ethical and legal thought be reconciled with facts about mental illness or brain dysfunction? Even now, we have evidence from imaging studies that our brains respond selectively to race (Pheips, 2001). Are we seeing activity related to social, or merely perceptual, judgment? Are differences innate or learned? Will the biologizing of the moral undermine its status as moral?

It is clear here that there are a multitude of questions and few answers. We may not even have a sense of what an answer to such questions would look like. Nevertheless, it is clear that as such questions are approached scientifically, the answers we get will shape our ethical views and, thus, will affect how we approach the ethics of neuroscience. As we learn more about the neuroscientific basis of ethical reasoning, as well as what underlies self-representation and self-awareness, we may revise our ethical concepts. This will then affect

how we evaluate the ethical implications of neuroscience for society. Similarly, engaging in ethical discussions about how to design and interpret neuroscientific experiments will affect what we can learn, control, and alter about the brain. The conceptual interconnections and feedback between the two main divisions of neuroethics are dense enough that it may be that distinctions between them can only be made roughly, and only in theory.

What's in a Name?

As with all newborns, picking a name is a difficult and contentious task. One of the most animated debates at the close of the NMTF conference was about the appropriateness of the label "Neuroethics." Some claimed it was an unfortunate name for this fledgling field, because ethics is the purview of philosophers, while the field clearly needs the concerted interaction of policy makers, lawyers, journalists, and the public, as well as the philosophers and neuroscientists. Others suggested that "Neuroethics" was ill-chosen because ethics excluded nonethicist philosophers and other humanists. I disagree on both counts. "Neuroethics" is a name well-chosen for a number of reasons. First, it is concise, catchy, and evocative. Second, it is a sad misconception of all too many that ethics is merely an academic exercise of philosophers. Rather, our ability to think and act ethically is arguably one of the defining things of what it is to be human: it is an inclusive rather than an exclusive term. Part of what it is to be a scientist, a doctor, a lawyer, a politician, or a journalist is to execute one's office in accordance with the values of one's profession and the society at large. Witness the Hippocratic oath, the courtroom oath, the swearing in before taking office, and the injunction not to fabricate stories or data. Ethics should therefore not be a domain foreign to nonethicist professionals. Moreover, in the time of Plato and Aristotle, it was considered imperative for every citizen to have a moral education and to take part in the ethical deliberations of society. It is perhaps reflective of some of the ills in our society that ethics is thought to be a philosopher's concern and not the common man's. But this is not a misconception we should yield to—it is an invitation to reeducate the public that ethics is a forum that needs the participation of everyone. Rather than capitulate to a narrow view of what ethics is and who it concerns, we should embrace the dialectical model of the NMTF meeting and demonstrate that ethics is as broad and inclusive a category as any.

We should not merely pay lip-service to this inclusiveness. Neuroethics has the potential to be an interdisciplinary field with wide-ranging effects. However, because it ultimately impinges on the well-being of the individual and our society, it is not a study that can or should be undertaken in the ivory tower. It is imperative that neuroethicists take part in a dialogue with the public. To make this possible, however, it is important in the short term to strive for "neuroliteracy" of the public and the media. We must make a concerted effort to make the subtleties of neuroscientific research accessible to the lay public via the media and refrain from the current practice of feeding it sound bites. For it is only with a nuanced understanding of the science, and a renewed trust in the goals of neuroscientists, that real progress will be made on these difficult issues. In the

last few months, we have heard just the first noises of such a dialogue. As Dana Foundation executive director Francis Harper aptly noted at the close of NMTF, "You can call it what you want, but the neuroethics train has left the station."

Selected Reading

Open your mind. (2002). *The Economist*. May 23.

Safire, W. (2002). The But-What-if Factor. *The New York Times*. May 16.

Fukuyama, F. (2002). *Our Posthuman Future: Consequences of the Biotechnology Revolution*. (New York: Farrar Straus & Giroux).

Damasio, A.R. (1995). *Descartes Error: Emotion, Reason, and the Human Brain*. (New York: Avon Books).

Greene, J.D., Sommerville, R.B., Nystrom, L.E., Darley, J.M., and Cohen, J.D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science* 293, 2105–2108.

Phelps, E.A. (2001). Faces and races in the brain. *Nat. Neurosci.* 4, 775–776.