CellPress

Neuron
NeuroView

An Ethics Toolbox for Neurotechnology

Martha J. Farah[1,*]
[1]Center for Neuroscience & Society, University of Pennsylvania, Philadelphia, PA 19104, USA
*Correspondence: mfarah@neuroethics.upenn.edu
http://dx.doi.org/10.1016/j.neuron.2015.03.038

Advances in neurotechnology will raise new ethical dilemmas, to which scientists and the rest of society must respond. Here I present a "toolbox" of concepts to help us analyze these issues and communicate with each other about them across differences of ethical intuition.

It is a truism that science is a double-edged sword. 20th century atomic physics revolutionized our understanding of material world and gave us new forms of energy but also created the deadliest weapons of all time, which continue to threaten civilization. The 21st century's most transformative science may well be neuroscience. We are living in a time of rapid progress, as neuroscientists gain new insights into the basic science of brain function and leverage them with a range of technologies from nanomaterials to machine learning. The articles in this issue of *Neuron* show the promise held by many of these methods for advancing basic science and treating neurological and psychiatric illness.

In the midst of this rapid progress, how can we encourage the development of ethical technologies and applications? Of course we will not have complete control over the field's development, and we will not even all agree on what constitutes an ethical use. Here I suggest that a constructive first step is to stock our ethics "toolbox." These tools will help us recognize ethical issues, analyze them, and communicate with each other about them.

Two Kinds of Tools: Consequentialist and Deontological

The ethics toolbox presented here has two main compartments, for consequentialist and deontological tools. Consequentialism is the ethical framework most closely associated with philosophers Jeremy Bentham and John Stuart Mill, according to which an act can be judged right or wrong depending on the expected value of its outcomes. Talk of "risk-benefit ratios," common in IRB (Internal Review Board) applications, is a use of consequentialist ethical reasoning.

This weighing of expected value is such a natural and obvious way of approaching ethical decisions in science and technology that it may seem pedantic to give it an "ism" name and cite its 19th century roots. Indeed, I recall a bioethics meeting at which a Nobel laureate scientist impatiently asked, "What's all this talk about? Just assess the benefits to patients against the risks and costs, and you'll know the right thing to do." But as this brilliant scientist came to appreciate, consequentialism alone does not fully accord with our ethical intuitions. For example, we would not be okay with sacrificing a healthy person to provide five lifesaving organ transplants, even though it seems right based on a simple calculation of aggregate benefits and costs.

The other widely used approach, which captures our sense of the wrongness of using a human being as an involuntary organ donor, is deontology, often associated with the 18th philosopher Immanuel Kant. The name "deontology" comes from the Greek word for "duty," and the approach determines what is ethical in relation to a set of moral principles that specify our duties and rights as persons. Our IRBs apply such principles as well as risk-benefit calculations. For example, even if risks are negligible and benefits are substantial, it would be a violation of a subject's right to autonomy to be enrolled as a research subject without informed consent.

Philosophers have attempted to reconcile the two approaches, for example, by considering the beneficial consequences of recognizing rights. This has never worked satisfactorily and so we are left with fundamentally different ethical systems. For many dilemmas the same decision is recommended regardless of which ethical system we use, but conflict can arise. Indeed, there are even cases in which different deontological principles conflict with one another or different ways of weighing consequences lead to different conclusions.

What this means for the toolbox offered here is that it cannot be applied algorithmically to reach a determinate answer. What it can do is capture and highlight morally relevant considerations in a given situation, to make more explicit the grounds for various ethical positions and to facilitate discussion when disagreement occurs.

The Deontology Compartment: Principles for Ethical Decision Making

Personhood. We all share an intuition that certain entities, including ourselves, are *persons* and hence have rights and duties, whereas others, including our furniture, are not and do not. These rights and duties are spelled out in the principles of deontological ethics. Many issues in bioethics have been analyzed in terms of personhood rights. For example, if a fetus is a person, then it has a right to life and abortion is wrong.

What is a person? For Kant personhood was related to the cognitive wherewithal (or cognitive potential, for the immature) to think and act morally. Others have used broader criteria, such as rationality and self-consciousness, but bioethics still lacks explicit criteria that capture our intuitions about who or what is a person without being circular (Farah and Heberlein, 2007).

Dignity. This concept was introduced into ethics by Kant as part of his explanation of how persons differ from objects. In Kant's terms, objects have prices, such that one thing can be fairly replaced by another when the prices are equal. This

is not true of persons; you would not entertain a trade for friends or family members regardless of the outstanding objective traits the proposed replacement has. Persons have what Kant called a "worth beyond value," which he termed *dignity*. Recently, this term has been used in a related sense by socially conservative bioethicists (Pellegrino et al., 2009) to encompass a kind of deep appreciation of humanity in all its imperfection and has thus figured in arguments against neurotechnological enhancement of humans.

Commodification. This concept refers to the extension of market value to parts of persons and their capabilities, including organs, reproductive capabilities, and cognitive capabilities.

Rights. These are moral entitlements, "must-haves" rather than "nice-to-haves"; in the words of the U.S. Declaration of Independence, "inalienable" from persons. An example is the *right to privacy.*

Beauchamp and Childress Principles of Bioethics. Bioethicists Thomas Beauchamp and James Childress crafted a set of specific principles to guide biomedical research and practice (Beauchamp and Childress, 2012). They are: *Respect for Autonomy*, which emphasizes the right to control our own lives, *Beneficence*, which refers to the duty to help others, *Nonmalfesience*, the duty to "do no harm," and *Justice*, which concerns broader duties to society, for example, promoting fairness and following the law.

Other Commonly Invoked Principles. The toolbox has many special-purpose tools, in the form of ethical principles that capture ethical intuitions in very specific contexts. Among these are the *wisdom of repugnance*, *natural is good*, and the *therapy-enhancement distinction,* which will be explained as they become relevant later.

The Consequentialism Compartment: Parsing Consequences for Ethical Decision Making

The basic tools of consequentialism are fewer and simpler than deontology (see Holland, 2003; Stanford Encyclopedia of Philosophy, 2014 to learn more about philosophical ethics and Farah, 2010 for an overview of the ethics of neuroscience). Here I will present a few concepts that are helpful in applying consequentialism to neurotechnology.

Kinds of Consequences. In Bentham's original *hedonistic consequentialism*, ethical actions are those that maximize everyone's pleasure. Because this seems to make the nucleus accumbens the arbiter of too much, a common variant is *preference or desire consequentialism*, where we act to maximize fulfillment of our more considered preferences. Of course our preferences themselves might be mistaken, so other approaches have been considered, including *perfectionist consequentialism*, which tells us to maximize the perfection or full flourishing of human potential.

Interests. These can be viewed as the consequentialist counterpart to rights, missing the obligatory nature of rights. They can be weighed relative to one another.

Externalities. Economists coined this term, referring to the effects of actions by one party on others who are not directly involved. This broadens the range of possible consequences that must be considered.

Sentience. To have interests, and thus figure in the consequentialist calculus, an entity must be *sentient*, that is, capable of experiencing perceptual and affective states. Humans are highly sentient, but at least some and perhaps all animals would also appear to be sentient.

Applying the Tools to Neurotechnology

Research Ethics

Human Subjects. A mix of consequentialist and deontological considerations guide our treatment of human subjects, including risk-benefit ratio and informed consent, the latter respecting subject *autonomy*. In research with neurological or neuropsychiatric patients, subjects may lack the competence needed for informed consent, and regulations then focus on protecting the person from harm, with *nonmalfesience* a particularly important principle.

Animal Subjects. The ethics of animal research is generally understood in *consequentialist* terms. Animals are viewed as *sentient* and we therefore strive to protect their *interests* as much as possible while accomplishing worthy research. The 3Rs of humane animal research (Russell and Burch, 1959) are a consequentialist amelioration of the ethical downside of animal research, based on a quantitative approach to degrees of goodness and badness. Animal research in neuroscience may be more ethically freighted than other fields, at least for modeling disorders of emotion and pain. Also relevant to the consequentialist calculation on the benefits side, the validity and usefulness of some animal models have been questioned (Nestler and Hyman, 2010). The idea of personhood and rights for some animals is an idea with some adherents (e.g., Regan, 1983).

Fetal and Embryonic Stem Cells. Those in favor of human fetal and embryonic stem cell research typically offer consequentialist arguments about the promise of these methods for curing disease. A deontological ethical analysis will depend mainly on whether fetuses and embryos are considered persons. If they are persons, then they have a *right* to life. Even if their fate would otherwise be the medical waste container, one would be *commodifying* them, or the reproductive functions of the parents, by using them. Although this is not my personal view, I think it is worth seeing that these objections arise from an approach to ethics that most of us have some sympathy for, even if we ultimately come down on the side of pursuing helpful new therapies.

Humanized Animals. Human genes and cells can be introduced into animal brains to create human disease models. Psychological changes can likely be induced by humanization; after all, behaviors can be transferred across nonhuman species (Balaban, 2005). Given how little we know about the likely psychology of nonhuman animals that have been humanized, it may be challenging to assess their levels of comfort, suffering, or frustration. This makes consequentialist analyses difficult to carry out. The primary deontological issue is which side of the person/nonperson line humanized animals are on. Sufficiently humanized primates might acquire mental capacities associated with highly developed sentience or even personhood (Greene et al., 2005). *Repugnance* is one motivator of ethical discussion of humanized animals, spurring us to question or, some might wish, limit the use of these methods.