Neuroethical Theories

MATTI HÄYRY

Practices and Questions

Neuroethics addresses moral, legal, and social questions created or highlighted by theoretical and practical developments in neuroscience. Practices in need of scrutiny currently include at least brain imaging with new techniques, chemical attempts to shift exceptional brain function toward normality, chemical attempts to enhance ordinary brain function beyond normality, and brain manipulation by other methods.[1]

Brain Imaging

Electrical activity recording (EEG), tomographic scanning (CAT, PET, etc.), and functional magnetic resonance imaging (fMRI) give information about the structure and dynamics of the brain both as it is without added external stimulation and as it works in experimental settings. Methods of measurement can be used clinically to ascertain brain death and coma, to diagnose epilepsy, and to detect abnormalities of the brain. Indirect information could also be produced concerning character traits, feelings of guilt, and the willingness to purchase goods and services.

Criminal investigators and lawyers are interested in lie detection by brain measurement, because this could, in difficult cases, give objective data that cannot be acquired by other methods. Techniques are also developed for revealing criminal intentions like terrorist tendencies, to be used before any harmful actions take place. Champions of neuromarketing are interested in applications that identify subconscious motives for buying products that the brain favors. More general brain mapping is believed to be useful in detecting a person's inclinations to mental health problems, racism, and violent behavior. Brain maps could, in the future, also reveal sexual preferences, people skills, neuroticism, risk aversion, pessimism, persistence, empathy, and intelligence in its many forms.

The main difficulties of brain measurement and imaging are currently their unreliability and openness to conflicting interpretations. An electrogram or image of the brain can tell that a person is lying or tends to feel other people's sorrows—or it can tell something else altogether, as uncontroversial scientific evidence is scarce. Courts of law have, in many countries, chosen to consider neuroscientific findings because of their appearance of objectivity and neutrality, which might be

The research for this paper was conducted as a part of *Neuroscience and Norms: Ethical and Legal Aspects of Norms in Neuroimaging (NeuroSCAN) / Ethical Concepts and Norms*, a project co-ordinated by Dr Tuija Takala and funded by the Academy of Finland (SA 1124638). An earlier version of my ideas presented here have been published in Finnish in Veikko Launis (ed.), *The Good and Bad of Neuroethics/ Neuroetiikan hyvä ja paha* (Kuopio: UNIpress, 2009).

Cambridge Quarterly of Healthcare Ethics (2010), **19**, 165–178.

doi:10.1017/S0963180109990430

problematic. Other questions are raised by the privacy, confidentiality, sensitivity, and eventual uses of the data pursued by public authorities, commercial actors, and others. How much are others entitled to know about us, how much are we obliged to know about ourselves, and why?

Medication of the Sick

Experts say that by measuring and imaging the brain—as well as by more traditional diagnostic methods—they can detect an individual's deviations from the average cognitive and affective skills of the population. Cognitive differences can be found in attention, concentration, and memory whereas affective diversity can be observed in proneness to depression, sleeping and eating disorders, and lack of libido.

When deviations are sufficiently pronounced, they can be labeled as illnesses and medication can be sought for them. Attention and concentration are boosted by methylphenidate (Ritalin) and amphetamine (Adderall). Memory aids are developed for people with dementia to enforce their ability to remember things and for people with post-traumatic stress disorder to selectively forget them. Fluoxetine (Prozac) removes depression or alleviates it, and many drugs that regulate brain function are researched for use in sleeping, staying awake, increasing appetite, decreasing it, and enhancing sexual drive.

Like any other novel drugs, neuromedications raise questions of safety, especially if they are used over long periods of time. In addition, there are the more specific challenges in medicalizing features that have previously been considered natural and in the impact that this can have on people's perceptions of themselves and others. Some individuals have always been absentminded, lively, forgetful, serious, gloomy, or in need of less sleep or sex. To call these conditions illnesses and diseases makes them undesirable in new ways. Consequently, many people want to be liberated from these features, but chemicals are not always the preferred solution: Individuals themselves often want to avoid the expected side effects, and behavior experts may recommend changes in life style, personal relations, work conditions, or other people's attitudes. There are also those who see the qualities they have as a part of their character and feel that they are being forced to alter their personalities. People who refuse drugs can be held responsible for, even guilty of, straying from the norms accepted by the community.

Medication of the Healthy

A line is arguably crossed when medication is employed to boost the cognitive and affective capacities of people who are already healthy by usual standards. Concentration, sharp memory, and good mood are assets in most jobs, and many people take drugs to enhance their performance in these respects in schools and work places although they have no medical reason to do so. The question is, is this somehow more questionable than curing recognized illnesses by pharmaceuticals?

It is not entirely clear why enhancements by neuromedication raise such high emotions. We can increase our alertness and persistence by drinking coffee or jogging regularly, but these activities are not frowned on in the same way. Memory and mood can also be improved by exercises that are openly advertised in the media and recommended in lifestyle guides. Why, then, do we have difficulties

with diligence and joviality when they are found in a pillbox? In all these cases, external means are used in order to regulate inner feelings and visible behavior.

The effects and side effects of neurodrugs provide easy targets for criticism. Prozac, for instance, is supposed to make miserable lives tolerable and tolerable lives fabulous, but reality is not that simple. The actual impacts of the drug can be more unpredictable, and the medication can lead to personality changes instead of the mere mood improvements that were intended. More generally, the safety of pharmaceuticals is always a concern.

Lack of safety is not, however, the only consideration that prompts negative responses to chemical neuroenhancement. Opponents can also claim that the practice is unnatural, unjust, and futile. Our character traits and abilities have their natural limitations, beyond which they should not be pushed except when a disease has to be prevented or cured. It is not right that some individuals can just by taking pills acquire talents that others have to work hard to achieve. And if everyone gets the same medication, the end result is that, as compared to others, everyone's qualities and skills remain the same—the most notable changes being that the bounds of normality have moved up and people have been exposed to the risks of medication for nothing. These are the kinds of objections that can be and have been leveled at attempts to improve healthy individuals by neurodrugs.

In the Head

In addition to medication, brain function can be affected by extra- and intracranial stimulation, neurosurgery, and human–machine interfaces. All these evoke various expectations and fears.

Extracranial brain stimulation by magnetic fields or electric current seems to provide relatively harmless ways to treat depression and, according to some studies, to boost creativity, lift mood, and strengthen cognitive skills. The use of intracranial electrodes is more dangerous, which is why deep stimulation of the brain is confined to the treatment of otherwise incurable and serious conditions like Parkinson's disease, epilepsy, depression, and obsessive–compulsive personality disorder.

Neurosurgery is employed to remove intracranial tumors and to treat brain trauma and artery diseases, severe pains, spasms, and epilepsy. Advances in imaging and microsurgery have made neurosurgical operations increasingly safer, but hands-on interventions on the brain continue to have their hazards.

Human–machine interfaces aim to transmit information from the outside world directly into our nervous system or from our nervous system directly to mechanical aids. An example of the former is the cochlear implant, which transforms sound waves into electric impulses and feeds them into auditory nerves, thereby making some deaf or near-deaf people sense sound. Reverse communication is required when prosthetic limbs or robotic appliances need to be guided by electric impulses of the brain. This latter type of technology is currently at an experimental level.

Interventions on the brain have an unpleasant ring to them partly because they have a long history of aggressive procedures, overambitious aims, and blatant mistakes. Electric shock therapy and cranial drilling evoke strong negative images despite the fact that they have had and continue to have an important role in neurotreatments. Human–machine interfaces do not always work even with cochlear implants, and the further development of people into cyborgs raises fears.

The invention of lobotomy earned António Egas Moniz the Nobel Prize for medicine, but in time the procedure was found to be not only inefficient but also dangerous to patients. Moniz's research was sloppy, erroneous, and (judged by current standards) unethical, and his reporting of the results possibly fraudulent, but the operation was enthusiastically embraced by the medical community in 1936 and it retained its popularity, suspicions notwithstanding, until the 1970s.

Philosophical Assumptions

Neuroethics is a field in which the strictest interpretations of the science that is assessed can be in conflict with the metaphysical assumptions of the methods by which the assessment is made. The sensible practice of neuroscience requires that the health, illness, and other features of the mind can be predictably altered by molding and regulating the structure and workings of the body (the brain). The sensible practice of practical ethics, on the other hand, requires that the behavior of the body (the embodied human being) can be changed by controlling the operations of the mind. If scientists claim that human actions are guided exclusively by the brain and ethicists claim that human actions are guided at least partly by the mind, then their positions are drastically incompatible.[2]

Mind and Matter

The relationship between mind and matter has for long been one of the most disputed issues in philosophy. In historical discussions the question was often framed either in terms of the existence and nature of an immortal soul or as a metaphysical debate on what sorts of entities the world is made of and how we can know about them.

Aristotle and Epicurus in Greek antiquity believed that the soul is an important but mortal part of our existence. Aristotle taught that the soul appears during fetal development and gives our body its form. Epicurus saw the soul as our thinking and sensing part that, like the rest of us, consists of material atoms. Both were convinced that souls perish with bodies at the end of life. Plato, their predecessor, held a different view. He reasoned that souls preexist bodies and live eternally after their demise. Aristotle regarded the soul as an immaterial entity; Epicurus and Plato defined it as very thin matter.

According to Christian philosophy, souls are spiritual and immortal whereas bodies are material and mortal. Because bodies are an endless source of distracting desire and lust, wise souls keep at a distance from them. At the dawn of modernity, René Descartes crystallized this dualism in the light of the anatomical knowledge of his time and surmized that (spiritual) souls and (material) bodies communicate through the pineal gland.

Post-Cartesian philosophers have developed many alternative answers to the questions of mind, matter, body, and soul. Idealists state that everything in the world is mental or spiritual, materialists maintain that everything is physical or material, and dualists allow the existence of two separate and fundamentally different substances. The general challenge for idealism is to explain what the things that we see as material are; for materialists, a similar problem is presented by our consciousness and thoughts, which at least appear to be distinct from matter. Our everyday thinking is most often dualistic: Senses seem to provide information

about the outside world, and internal beliefs and acts of will seem to guide our external behavior.

Thinking and the Brain

Neuroscientists believe that brain scans and imaging measure thinking and character traits and that medication and surgery can change them. But what exactly is measured and changed? Is it really thinking itself or is it only its representation or counterpart in the brain?

Alan Turing, a pioneer in computer technology after World War II, was convinced that he was developing a thinking machine, and he suggested a test, now bearing his name, which would check what the electronic brain can do. A person exchanges messages with someone or something else—the second party is not visible to the first one and can be either another person or a machine programmed to converse like a human being. If test subjects are convinced that they communicate with another human being when they actually communicate with a machine, the device passes Turing's test and it is by his definition "thinking."[3]

The thought experiment gives rise to a number of considerations. One of these is that the beliefs of the test subjects do not, of course, change reality. People can think that they are in contact with a human being, but this does not change the fact that their partner in communication is an inanimate object. If the machine is nevertheless described as "thinking," this quality is intersubjectively defined and subjectively perceived rather than objectively apprehensible. This line of reasoning can also be applied to mutual human interactions. When we talk with people, are we in touch with their minds ("persons") or only with their brains ("machines")? If the latter is true, do minds exist at all? Or if the former is normally true but the particular person we are talking to is on heavy neuromedication, are we addressing a front produced by pharmaceuticals and is the actual (unmedicated) person hiding behind the manipulated brain?

If we hold on to the idea that minds exist separately from bodies, neuroscience is not directly involved with thoughts and character traits—instead, it measures and controls their counterparts in the brain. This does not necessarily mean that mental processes guide material ones in ways disputed by science. The doctrine of epiphenomenalism claims that dualism is true, but that the transactions between body and mind are one-directional. Changes occurring in the brain are registered and experienced in the consciousness, but consciousness cannot influence the brain or our behavior. Persons as subjects of self-awareness are "ghosts in the machine"—bystanders of their own existence who cannot steer the course of their lives but who, because of illusions created by their brains, imagine that they can. People who have been subjected to neuromedication sometimes report that they have been simultaneously aware of their medicated *and* unmedicated selves. In epiphenomenalist terms, this means that the brain sends the consciousness two different strings of messages, both of which it senses.

Freedom and Determination

Most Western philosophies assume that human actions are initiated by our own free choices. The movement of animals is guided by external incentives and inborn instincts, but people can modify their behavior by deliberation and self-discipline.

The idea is natural enough, but it is not obviously compatible with the theological and scientific beliefs of the West.

Europe is traditionally a Christian continent (as are other regions influenced by European culture), and, in Christianity, God has been seen as omniscient and omnipotent. On closer look, however, the ideas of omniscience and free will are probably contradictory. If God knows at each moment what a person is going to do next, the person does not have any genuine power to choose among alternatives—what will be done, will be done. Theologians and philosophers have been aware of this problem for centuries, and several solutions have been suggested. Augustine of Hippo taught that only the first two people, Adam and Eve, had free will, a privilege that was removed from their descendants after the crucial choice of the forbidden fruit. Later attempts to reconcile the tension have included speculations on a weaker God who originally created the universe but is not fully in control of its workings any more.

The development of natural sciences created a similar problem for secular thinking. In the Newtonian worldview, all events are preceded by events that have caused them, which, in their turn, are preceded by events that caused them, and so on. Turned around to point to the future, this model implies that everything that is ever going to happen has already been set in motion and determined: New elements like acts of the will cannot be introduced to change the course of history. Post-Newtonian quantum mechanics has challenged the idea of complete determination, but indeterminism leaves no room for free will either. Unless our choices coincide with quantum leaps, we are left as powerless in a model that attributes the occurrence of some events to pure chance as we are as prisoners of the inexorable laws of physics.

This tension has also been recognized in philosophical traditions, and various responses have been proposed. The Stoic solution is to accept the inevitability of external events and their causation and to concentrate on controlling one's inner moods and attitudes. The world is what it is; those who are wise acknowledge this fact and those who are not waste their time and energy by trying to change immutable chains of events. The Epicurean answer was to find a niche for the free will in the sometimes random movements of atoms—this is the line of thought continued by physicists who today seek God and autonomy between quarks. Another popular reaction is to separate how things are and what we can know about them. Free will means, according to this model, simply that we do not yet know what natural causation will make us do next.

Bridge to Ethics

Immanuel Kant was well aware of the problems of mind, matter, determinism, and free will, especially in the context of ethics. Individuals cannot be responsible for things that they could not have caused or prevented. Pure logic dictates that there cannot be any moral ought to do what we cannot do. Accordingly, if mind cannot control matter and our will does not guide our actions, morals and ethics cannot exist in any genuinely action-guiding sense.

Kant solved this puzzle in his philosophy by dividing the world into two parts and by naming different modes of existence for them. The externally observable realm of physics follows the laws of determination, and our knowledge of it is filtered by the human categories of space, time, and causality. Brain events

studied by neuroscience belong to this realm dictated by natural necessity. The internally experienced world of morality, on the other hand, follows the logic of human reason, and its operations are guided by the moral law, knowable to us only by philosophical reflection. Rules concerning research and its implications, studied by neuroethics, belong to this sphere of human freedom.

Unfortunately, Kant could not explain the relationship between his two worlds any better than other dualists before or after him. Freedom is a necessary condition of morality, so it has to be assumed to have a foundation for philosophical studies in ethics. But the way the free will works and affects events in the empirical reality remains unexplained. Kant himself conceded that action according to the moral law in a world of causal necessity is a great mystery.[4]

Mystery or not, the fundamental tenets of neuroscience and neuroethics become compatible only by assuming something along the lines that Kant suggested. The epiphenomenalism favored by scientists leaves room for the view that the mind (consciousness) is in some indefinable way informed about the actions of the body (brain). But the freedom of choice presupposed by normative ethics requires that the mind can also will things and by its actions somehow influence the brain and eventually affect external behavior. If this notion is unsatisfactory, we have to think that the entire question is incorrectly posed and that what we are dealing with here is a pseudo paradox created by our sloppy use of language.[5]

Antitheories

In neuroethics, like in applied ethics more generally, one methodological choice is to claim that moral theory is not needed in the field. Problems can be solved on a case-by-case basis by employing intuitions or common sense or by applying widely accepted principles, largely ignoring their theoretical underpinnings. These antitheoretical models can provide answers for relatively simple ethical questions, although their use will sooner or later lead to conflicts and ambiguities that can sometimes be clarified by an understanding of prevailing moral doctrines.

One Case at a Time

Contextualism and casuistry state that issues can be resolved by becoming acquainted with the details of the case, listening to all the parties involved, and comparing the situation to instances in which we would more self-evidently know how to act. In considering, say, the need for informed consent in brain imaging, the correct response will be found by letting the particulars of each case speak for themselves.[6]

When the number of people to be scanned or treated increases, it will become more difficult for every researcher and healthcare professional to examine each situation on its own terms. This is why ethicists and regulators endeavor to find more general rules to be observed, at least in noncontroversial circumstances. The description of the practices (in the first section of this article) shows that neuroscience creates challenges related to safety, accountability, privacy, consent, trust, confidentiality, responsibility, stigmatization, and discrimination. Moral experts and public authorities can recognize these categories but also keep an eye on contextual variations.

The safety of experiments, measurements, and procedures is assessed by calculating their risks and by making them commensurate with expected benefits

and harms. Concrete findings are usually more reliable than ones that need theoretical interpretation: Brain tumors are easier to see in magnetic resonance images than character faults. If patients are in serious and immediate danger, operations can be justified despite uncertainties and risks, but when situations are less urgent, invasive procedures cannot be recommended with similar intensity.

If the information obtained by scanning and imaging is sensitive and personal and thereby belongs to a person's sphere of privacy, the dissemination of data can be harmful even if no (other) detrimental consequences ensue. Sharing the information with researchers and medical professionals, as well as more concrete risk taking, can be legitimized by the free, uncoerced, and informed consent of the patient or research subject. But this presupposes that the objects of tests and experimentation trust neuroscientists and that neuroscientists respect the confidentiality of the data they have in their possession. Knowledge in the wrong hands can, depending on the case, cause violations of privacy, futile accusations, stigmatization, and discrimination.

American Principles

Those who have wanted to clarify ethical thinking without going too deep into philosophical theory have proposed that morality can be encapsulated in a handful of principles capable of guiding practices and their assessment. The best known formulation of this idea was first presented in 1979 by Tom Beauchamp and James Childress, who suggested that most bioethical questions can be answered by applying the principles of autonomy, nonmaleficence, beneficence, and justice.[7]

Respect for autonomy decrees that individuals must be allowed to choose their participation in treatments and experiments freely, informedly, and without coercion or duress. In neuroscience, these requirements have to be interpreted rather loosely to retain a semblance of voluntariness in imaging, medication, and surgery. The ambiguities in measurements, the side effects of drugs, and the risks of operations are, in many cases, far too sizable to be accepted by healthy volunteers knowingly and without compensation. Human guinea pigs are, subsequently, recruited from among patients whose conditions force them to take part in experimental treatments, ordinary citizens whose consent is uninformed and questionable, and researchers and students whose pressure to participate comes from their work environment and reward in the form of study materials.

Neuroscience is not always safe, and its beneficial effects can be uncertain, but in the "American principles" model introduced by Beauchamp and Childress, autonomy and consent usually justify even dangerous and futile practices. The responsibility is transferred to the individuals giving the permission, and, especially if they are confronted with otherwise incurable diseases, it is easy to accept even considerable risks. To protect people from hazards that they want to choose on good personal grounds is seen as inexcusable paternalism.

Justice has, within this model, at least two dimensions that are related to neuroethics. The first is that allocating resources to experimental research and therapies is dubious, because more tangible health benefits could be achieved by preventive medicine, social reforms, and tested treatments not currently available to everyone. The second is that people should perhaps volunteer for medical trials more actively, because otherwise they will, with advances in science, eventually enjoy the benefits of activities that they have hindered by their indolence.

European Values

The principles defined by Beauchamp and Childress have been criticized for emphasizing pragmatic over moral considerations and individualism over communality. Attempts have been made to replace them, for instance, by the "European values" of dignity, precaution, and solidarity.[8]

Dignity can, depending on the interpretation, mean several things: the sanctity of every human life, the moral self-determination of individuals, dignified treatment, and so on. Neuroscience can offend some of these expressions of human worth by its very presuppositions: Immortal and immaterial souls do not exist, the freedom of the will is an illusion, and humanity is not fundamentally different from the rest of nature. And even if these affronts are overlooked, the problems that remain include violations of privacy and discrimination based on physical diversity. Many of us can prefer to keep the information produced by brain imaging to ourselves, and it is not too farfetched to believe that suspicions of a tendency to violence can lead to unequal treatment regardless of our actual behavior toward others.

The principle of precaution, first employed in environmental contexts and then extended to other areas, states that if a new technology can be hazardous, it should not be implemented before its safety has been established. The point is to shift the burden of proof from the opponents of novel practices to their proponents. A conceptual difficulty in the application of the principle is that new technologies are often by definition potentially dangerous simply because the consequences of untried solutions are routinely unknown. It stands to reason, however, that original ideas should sometimes be tested if they promise to remove prevailing ills. Interpreted literally, precaution would require neuroscientists to prove in advance that what they plan to do is harmless.

Solidarity demands that individualistic concerns for justice be replaced or complemented by community thinking and the protection of the vulnerable. Individual consent is not enough to safeguard this value. The competing model claims that a person's own permission and preference to be tested or medicated provides a sufficient warrant. The advocates of solidarity stress that people's preferences can be manipulated by false hopes and social pressures. Excessive fear of wrongful paternalism can lead to negligence when patients or experimental subjects act under duress and without proper understanding of the risks.

The Need for Background Theory

All the values and principles presented in the antitheory approaches can be helpful in ethical deliberations, but none of them can give us universally accepted rules and norms in real-life situations. This is prevented by ambiguity of concepts, conflicts within models, and disputes between them.

Every term used in moral evaluations can be given many different readings. Consent, for instance, can be required in at least four distinct senses. *Actual consent* means the patient's or research subject's own permission to procedures or experiments. *Assumed consent* is in use when the persons themselves cannot be asked because of unconsciousness or urgency, but their prior opinions and behavior are thought to imply that they would authorize the treatments. *Hypothetical consent* is a philosophical construction that postulates that, as good citizens or rational agents, individuals are bound to acts as physicians, scientists, philosophers,

or public decisionmakers wish. *Proxy consent*, finally, does not require individuals themselves to agree with the practice at all; instead, their family members or representatives give the go-ahead. As reasonable people can disagree on the type of consent that should be required for neuroscientific work, it is unrealistic to believe that context and circumstances could unequivocally determine which interpretation must be adopted.

Within each model—casuistry, contextualism, and principlism—focus on different details or rules can lead to different ethical recommendations. From the viewpoint of healthcare professionals, respect for autonomy and the avoidance of harm would probably clash in the case of a client who asks for depressing neurodrugs in order to write an existentialist novel.

''American'' and ''European'' values can in practical situations be in conflict with each other. One model states that competent adults can decide what they subject themselves to; the other maintains that dignity, solidarity, and precaution can define additional or alternative conditions to this formula. Because the parties cannot in all cases come to mutual agreement, it is futile to wait for universal solutions from these quarters.

None of the difficulties stated here is insurmountable to ardent proponents of their own version of antitheory ethics. They can choose the meanings of terms so as to fit their thinking, solve conflicts between principles according their own intuitions, name one of the principles or values (the commonest choices are autonomy and dignity) as paramount, and declare their construction as the most rational or most moral. Ethical edifices like this are, however, seldom widely (and never universally) accepted, and this is why many philosophers have thought that a deeper moral background theory could provide a firmer basis for our responses.

Three Theories

The main theories of normative ethics link the rightness of actions either with their physical and mental consequences, their moral reasonability, or their concordance with human nature. When neuroscientific research, its application, and the regulation of the two are considered in the light of these general systems, a philosophical setting emerges for their comparison and evaluation. In the final analysis, however, the real obstacle is the choice of just one theory out of the three candidates. On what grounds could one of them be regarded as better than the others?

The Consequences of Choices

According to the moral doctrine of utilitarianism, a choice (to act or not to act) is the right one if and only if its concrete and measurable consequences to human and other sentient beings are at least as good as those of any of its alternatives open to the decisionmaker at the time. In theory, agents should take into account all the consequences of all possible options, including their impacts on other people's feelings as well as their indirect and remote repercussions. In practice, however, utilitarians who favor scientific advances confine their attention to relatively immediate harms and optimistically estimated benefits, excluding from their calculations offended feelings, popular fears, and moral anxieties.[9]

Neuroscientific activities can be assessed easily by using this approach. Everything that poses a risk of harm to patients or experimental subjects is

forbidden unless it can be independently justified by consent or urgency. On the other hand, everything that can be justified by consent or urgency is permitted as long as some benefits can be expected for the individuals in question or for others. Brain imaging and external manipulation are usually seen as harmless and, subsequently, unproblematic. Invasive brain operations are dangerous, but can be justifiable when nothing else can be done to help patients. Medication of the sick and the healthy is equally allowed, provided that safety issues are under control and official authorizations are in place.

Utilitarianism offers a philosophical foundation to the ''American'' principles of bioethics—like the model's inventors intended it to do. Generally speaking, utilitarianism forbids and permits the same neuroscientific experiments and applications as Beauchamp and Childress's doctrine, but, in addition, its proponents can claim that they can resolve conflicts between principles better than advocates of principlism. If autonomy and nonmaleficence clash, the expected impacts on both of them are converted into compatible units in terms of human well-being. If restrictions of autonomy cause physical or mental ill-being, this effect has to be compared with the harm probably inflicted by the medication. The alternative that carries the lowest net disadvantage and the highest net advantage ought to be chosen, regardless of possible violations of other principles.

Moral Reasonability

According to Kantian ethics, we should act only by rules that rational agents could accept as universal laws for themselves and for all other rational agents. Kant's original idea was essentially formal, or logical. Because morality presupposes freedom of choice and because we can be free from natural causation only in the realm of intellect and rationality, we must, in moral matters, operate exclusively in terms of reason. Furthermore, because reason is the same for everyone, the rules it dictates are also the same for everyone. As seen in the discussion concerning free will, however, the method by which freedom and reason guide our actions is a puzzle—a mystery that lies at the core of our being but cannot be fully explained. This is why Kant, like many of his followers, saw the protection of humanity as he understood it as the best strategy in uncertain cases. Although Kantian ethics, taken literally, applies only to rational agents and their voluntary choices, the role of morality can, in less literal versions, be extended to the defense of humanness—as an essence or as an idea—even over and against the expressed wishes of the individuals involved.

If morality requires only the acceptance of reasonable people and if the sphere of reasonable concern is limited to consent and tangible harm, neuroscience is viewed as favorably by Kantianism as it is by utilitarianism. Physically dangerous procedures like deep stimulation of the brain are prohibited unless they are the last resort for saving a patient, whereas lesser risks can be left for patients to assess. This theory, like utilitarianism, is perfectly compatible with the ''American'' values of Beauchamp and Childress, although in conflict situations it is more prone to support autonomy than the avoidance of harm.

If, on the other hand, morality requires the protection of humanity against violations and changes, the picture becomes radically different. Jürgen Habermas has argued that the continuity of human life, at least as it is understood in Europe, presupposes that all present and future individuals are allowed to participate in

making decisions that concern themselves.[10] Neuroscientific applications like chemical enhancements beyond health and advanced brain–machine interfaces can be construed as attempts to create new people by technological means. Personalities can be drastically altered by medication, and self-images can be irreversibly transformed by machines to which people are attached. If either of these occurs, original decisionmakers who have given their consent to neuro-improvements arguably cease to exist and are replaced by new people who have not been able to take part in the choice. Habermas believes that this would violate humanity in the Kantian sense and should therefore be forbidden.

Concordance with Human Nature

Ethical thinking based on Aristotle's philosophy sees people as social animals whose main goals in life are self-preservation, reproduction, and the pursuit of truth. As neuroscience aims at increased understanding and its applications at the protection or restoration of health, neuroresearch and therapies go well together with Aristotelian ethics. The enhancement of cognitive and affective abilities beyond normality poses questions in this context, though.

Michael Sandel has criticized the current culture of manufacturing and molding children in affluent societies and claimed that it gradually undermines solidarity between citizens. Sandel insists that children should be seen as gifts, and that the determination of their abilities and qualities should be kept beyond parental control. People who understand the ''given'' nature of our lives and fortunes react positively to the ideas and practices of risk sharing and joint responsibility for offspring, dependents, and vulnerable individuals and groups. Human lives are full of unintended twists and turns that we cannot predict. The gene manipulation and neuromedication of children create false views on how much we can and should control each other's features and behavior. In this atmosphere, parents who have *not* enhanced their progeny's alertness and attention by Ritalin and Adderall can be deemed guilty of their children's learning difficulties and social problems. Individual choices (parental and otherwise) are emphasized, the communal nature of humanity is forgotten, and the foundation of solidarity crumbles.[11]

Sandel and Habermas reach similar conclusions in their evaluations of scientific advances: Both believe that the chemical and technological enhancement of individuals paves the way to the end of humanity as we know it. Sandel wants to make clear, however, that his reasons for this verdict are different from his rival's. He argues that conceptual stress on consent and autonomy keeps Habermas firmly in the liberal and individualistic camps, which are opposed to the ideals of community and shared responsibility. Sandel himself sees the future of humankind in communal traditions and believes that people can properly deal with the ''externally given'' elements of their lives only by recognizing their existence and by warding off their effects in a spirit of solidarity.

The Wisest Choice?

Traditional theories of normative ethics yield different results on the morality of neuroscientific research and its applications. All of them have their advantages and their disadvantages, and none of them work for everybody in all circumstances.

Knowledge and understanding of the theories can, nonetheless, facilitate sensible decisionmaking on several levels.

Utilitarianism offers a down-to-earth and easily applicable model for solving ethical issues. It seems clear that, when decisions are made in science and healthcare, the impacts of these decisions on people's well-being ought to be taken into account. But the main tenets of Kantianism are equally attractive. We should, no doubt, use our reason and respect everybody's humanity when we make our choices. And similar considerations can be extended to Aristotelian thinking. The basic ingredients of human life, including its communality, should certainly be recognized and taken on board in research and medicine.

The problems of utilitarianism are linked with the assessment of consequences and with the rejection of values and rules not related to physical and mental contentment. We can never know all the outcomes of our actions. Focusing on only some of them raises questions of definition and demarcation, and even if all consequences favor a particular choice, people can turn it down if it is seen as unjust or immoral. The difficulties of Kantianism are based on its formality and its reliance on human dignity. It is tempting to think that moral rules have to be reasonable and universal, but this general thesis does not without specifications tell us what to do in practical situations. The specifications, in their turn, bring about disagreement: Respect for humanity in the Habermasian sense can work for some but it does not work for all. And the Aristotelian model has issues associated with explaining the normativity of human nature and the primacy of community. We are self-preserving and reproducing seekers of the truth, but we are also, and sometimes simultaneously, indifferent exploiters of natural resources and ignorers of each other's interests: Why should some of these features but not others determine us as moral beings? Stress on communal life and traditions gives rise to another kind of potential discrepancy: Not all communities advance the welfare of their members, and when they do not, why should they be preserved and promoted?

Despite these concerns, awareness of the main theories of ethics and continued discussion on their merits and demerits can offer some guidance in the sensible control of neuroscientific research and its applications. Without knowledge of theoretical moral thinking, those who study and interpret laws, give regulative recommendations, and pass legal verdicts can overlook significant aspects of human thinking. And although the attitudes of legislators are always ultimately based on ideologies and political convenience, some understanding of the major ethical traditions would not necessarily go amiss even in their case.

Notes

1. A useful overview of the dimensions and questions of neuroethics is provided by the University of Pennsylvania web site *Penn Neuroethics Briefing* accessible at http://www.neuroethics.upenn.edu/ (last accessed 3 Aug 2009).
2. Comprehensive views on the philosophy and ethics of neuroscience have been presented in, e.g., Walter H. *Neurophilosophy of Free Will: From Libertarian Illusions to a Concept of Natural Autonomy,* C. Klohr, translator. Cambridge, MA: MIT Press; 2001; Rees D, Rose S, eds., *New Brain Science: Perils and Prospects*. Cambridge, UK: Cambridge University Press; 2004; Gazzaniga MS. *The Ethical Brain*. Washington, DC: Dana Press; 2005; Glannon W. *Bioethics and the Brain.* New York: Oxford University Press; 2007.
3. Turing A. Computing machinery and intelligence. *Mind* 1950;59:433–60.

4. See, e.g., Häyry M. The tension between self-governance and absolute inner worth in Kant's moral philosophy. *Journal of Medical Ethics* 2005;31:645–7.
5. Tom Buller addresses the question of scientific determinism versus free will in more detail in his contribution to this issue.
6. See, e.g., Timmons M. *Morality Without Foundations: A Defense of Ethical Contextualism.* New York: Oxford University Press; 1999; Jonsen A, Toulmin S. *The Abuse of Casuistry: A History of Moral Reasoning.* Berkeley, CA: University of California Press; 1988.
7. Beauchamp TL, Childress JF. *Principles of Biomedical Ethics*, 6th ed. New York: Oxford University Press; 2009.
8. See, e.g., Häyry M. European values in bioethics: Why, what, and how to be used? *Theoretical Medicine and Bioethics* 2003;24:199–214; Takala T, Herissone-Kelly P, Holm S, eds. *Cutting through the Surface: Philosophical Approaches to Bioethics.* Amsterdam: Rodopi; 2009, esp. chaps. 8–12.
9. See, e.g., Häyry M, Utilitarianism and bioethics. In: Ashcroft R, Dawson A, Draper H, McMillan J, eds. *Principles of Health Care Ethics*, 2nd ed. Chichester: John Wiley & Sons; 2007:57–64.
10. Habermas J. *The Future of Human Nature*, Rehg W, Pensky M, Beister H, trans. Cambridge, UK: Polity Press; 2003.
11. Sandel MJ. *The Case against Perfection: Ethics in the Age of Genetic Engineering.* Cambridge, MA: The Belknap Press of Harvard University Press; 2007.