

Neuroethics: what else is new?

Alex Mauron

Institut d'éthique biomédicale, Centre médical universitaire, Genève, Switzerland

Funding / conflict of interest: No funding. No conflict of interest.

Summary

Bioethics has changed considerably over the years, the increasing diversity of topics of bioethical interest having led to specialisation if not fragmentation. In this context the term “neuroethics” emerged, and it soon became clear that it could be understood in two ways. The “ethics of neuroscience” assesses the ethical and philosophical implications of research in neuroscience and of the application of these discoveries, especially as regards enhancement of “normal” human faculties or new challenges to privacy. The “neuroscience of ethics” basically turns the tables on ethics: it is ethical reasoning itself that becomes the object of neuroscience research. This provides a “behind the scenes” glance into moral reasoning, challenging aspects of moral philosophy and of the conventional philosophical interpretation of human action and free will. On the more practical side, neuroethics increases the level of awareness of the relatively neglected area of diseases of the brain-mind.

Key word: neuroethics; philosophy of mind; free will; mental health

Ethics as a suffix

Some time ago, we lived in a world of ethics, period. It was inculcated by family and school, thundered from the pulpit in church, and was expertly concocted and minutely dissected in faculties of theology and philosophy departments. In addition to this supposedly universal ethics, there were specialised moral codes for specific professions, such as physicians and lawyers, but these were largely considered an internal matter for these professionals themselves and had relatively little to do with general ethical norms or theories. Change came primarily from medicine, and later also from biotechnology. The therapeutic revolution of the 20th century gave an enormous boost to the healing power of the physician. Paradoxically, it also increased the power of the patient who, in this new world where medical care and medical choices actually made a difference, increasingly asserted his right to a say in these choices [1]. Furthermore, the ethics of medical experimentation on human subjects became a major moral concern in the wake of the discov-

ery of Nazi and Japanese atrocities perpetrated in the name of “medical research” before and during World War II; but it also arose from the revelation that even in democratic countries there were widespread medical research practices that had scant regard for the rights of medical research subjects. Finally there were the novel moral questions raised by genetics and reproductive technologies. Around 1970, ethics received a “bio-” prefix and the result was a social and political movement which in turn resulted in the new, now well established academic discipline of bioethics. Some did ask at the time (and occasionally still ask today) why a new academic field was needed to do what they saw as the mere application of ethical concepts to a set of particular issues. One plausible answer lies in the retroactive effect of bioethics upon ethical theorising itself. Academic philosophers inhabiting the dizzy heights of metaethical speculation were suddenly welcomed back into the real world, and as a result bioethics changed the way in which some philosophers and some theologians reflected on ethics in general [2].

From bioethics to neuroethics

In the last two decades the number of topics explored by bioethicists has exploded and there is a constant temptation to further fragment the discipline. For instance, the increased power of genetic diagnostic tools, the complete sequencing of the human genome and the prospect of high-throughput sequencing of individual human genomes apparently call for a specialised field within bioethics. Indeed the term “genethics” did appear in a number of articles and book titles during the 1990s and sometimes later. In the end, however, the new term failed to fly [3]. It was recognised that the *modus operandi* of bioethical analysis was generally appropriate as a means to reflect on the new issues raised by human genetics and genomics, no matter how diverse and challenging. Might not the term “neuroethics” be considered redundant on similar grounds? Clearly, there are many issues within the scope of neuroethics which have strong similarities with ethical dilemmas in other areas of bioethics. The classical case is the question of how to deal with incidental findings in neuroimaging procedures such as fMRI. The increased power and resolution of these investigative and diagnostic methods are bound to make “unwanted” incidental discoveries about a patient or research subject increasingly frequent [4–5]. This in turn raises difficult questions about disclosure to the patient and/or research subject. Yet these issues are in a large measure sim-

Correspondence:
Prof. Alex Mauron
Institut d'éthique biomédicale
Centre médical universitaire
1, rue Michel-Servet
CH-1211 Genève 4
alexandre.mauron@unige.ch

ilar to those raised by genetic diagnostics, especially multiplex screening for many genes at one time. The dilemma can even be formulated in very general terms: patients and doctors have a desire and a need for biomedical progress in diagnosis and therapy. As a result, however, they will have to live with the fact that progress in diagnostics will increasingly give answers to unasked questions as well as to those that have been asked. How to handle this situation within the fiduciary relationship of patient and physician is perplexing enough. If in addition one takes on board the interests of third parties who might be interested in knowing about the presymptomatic status of persons destined to develop a debilitating illness, the problem becomes even more troubling. For example, it is significant that ethical dilemmas about predictive testing for Alzheimer's disease raise similar questions, whatever the nature of the predictive data considered, such as genetic or neuroimaging data [6–7]. “Genethics” and “neuroethics” coincide in this case.

The ethics of neuroscience

Some neuroethical problems, however, do have a stronger claim to be uniquely “neuro” because they are specifically relevant to the causal nexus uniting brain and mind. Ingesting particular substances that affect brain physiology with a view to altering mind-states is probably a practice as old as humanity itself. Nevertheless, psychopharmacology is attracting specific neuroethical attention today. Interestingly, the focus here is less on the therapeutic use of psychoactive drugs (a controversial matter in the heyday of the anti-psychiatry movement in the 1960s), than on the enhancement of mental faculties, especially cognition, by drugs. In one sense this area of ethical reflection still has a rather anecdotal factual basis. At present there are a mere handful of drugs (methylphenidate, modafinil, SSRI antidepressants) that attract a huge volume of ethical commentary about their use for enhancement i.e., beyond their standard therapeutic indications [8]. Nevertheless, these practices tie in with intricate philosophical problems surrounding the robustness of the therapy/enhancement distinction as well as the notion of an authentic self that may or may not be altered by enhancement practices. The underlying intuition is of course that altering the brain-mind is in some sense more radical and more basic than to alter some “outer”, superficial, or accidental characteristic of a person. But then the problem may reside with our intuitions, which may well be misleading in contexts that are completely novel for mankind. Somewhat similar discussions relate to non-pharmacological interventions on the brain such as transcranial magnetic stimulation or deep brain stimulation. The latter, however, is much more invasive than the former and may not even be considered for non-therapeutic uses.

Neuroimaging (especially fMRI, which is better adapted to non-medical research than other brain scanning techniques) also titillates our philosophical curiosity, not by creating radically new philosophical questions but by bringing classical ones into sharper focus. Issues such as the subjective/objective distinction and the private nature of our unexpressed thoughts were and still are philosophi-

cally perplexing. Indeed, no matter how often philosophers and neuroscientists have deconstructed Cartesian dualism, most of us still experience our consciousness of the world as a “Cartesian theatre”, where a homuncular “I” scans the stream of thoughts coming from other thoughts or from the “outside” [9]. Most importantly, in the Cartesian theatre, only “I” have a valid ticket to watch the show, which underscores the privacy and inviolability of “my” inner life. But presently these immemorial intuitions may face a serious challenge from brain imaging. Journalistic and commercial hype about fMRI-based “mind-reading” or lie detection notwithstanding, it is the case that fMRI provides access to some aspects of mental life that were previously unknowable. For instance, such studies are increasingly showing robust correlations between specific brain activation patterns and the subject's viewing different images such as animals, human faces or non-significant patterns [10], and researchers are tackling the challenge of actually reconstructing images viewed by subjects from fMRI recordings [11]. These approaches are not limited to decoding of visual data, but also investigate more abstract tasks such as (not) recognising words in a language that one is (not) familiar with [12], or exploring the neural correlates of initiating and reflecting on one's actions [13] and other-regarding attitudes such as empathy [14]. If and when these results are confirmed by more research and eventually move from the research laboratory to less peaceful places such as the courtroom or the police station, the ethical challenges are pretty obvious. The defence of privacy and confidentiality, and the protection against self-incrimination will need a great deal of preventive ethical reflection. Of course, privacy and confidentiality constitute ethical issues in their own right in many bioethical contexts, such as sexually transmitted diseases, false paternity, addiction and the like, but neuroimaging may one day provide access to private information at its very source in the human brain, bypassing the selection and self-censorship inherent in human speech. This could be the ultimate self-incrimination: cutting out the middleman, which in this case is yourself!

The neuroscience of ethics

We have seen that neuroethics has some claim to be special in the same way that bioethics became special around 1970, because in both cases ethical work on particular concrete topics feeds back to more basic philosophical issues. This impression is even stronger when we consider the retroaction of neuroethical analysis not only on the philosophy of mind, but on philosophical ethics itself. As Adina Roskies noted early on, neuroethics is poised to become not only the ethics of neuroscience but also the neuroscience of ethics [15]. In fact, neuroscience is turning the tables on ethics in more ways than one. Firstly, ethical reasoning has long been an object of psychological research, but the tools of current neuroscience are opening new vistas for such studies. In the last decade a number of neuroimaging studies have been done on subjects reflecting on an ethical dilemma, typically one of the many variants of the classical “trolley problem”, which can be stated thus:

- A An out-of-control trolley is hurtling down a track where five workers are absorbed by their task some distance downhill, unaware of the danger to their lives. You can flip a switch that will cause the trolley to turn onto another track, where there is only one worker, thereby causing (?) or allowing (?) one person to be killed instead of five. Do you flip the switch?
- B The story is changed somewhat: you are standing on a footbridge above the tracks, next to a very fat man whom you could push onto the track. While he would stop the trolley on account of his enormous bulk, he would be killed while the five workers would be saved.

Extensive psychological research has documented what most people's moral intuitions are: roughly speaking, they would flip the switch in version A of the story and not push the fat man in version B. In other words, they would be utilitarians in A (they compute the outcome of their choice in terms of the number of lives lost or saved), deontologists in B (they apply the general prohibition on intentional, direct killing). Furthermore, fMRI studies have documented several brain areas involved in making such moral judgments [16], complementing data from pathology which suggest that lesions to the ventromedial frontal cortex alter the "normal" pattern of responses in moral evaluations [17]. These and more recent studies are providing a finer-grained picture of the various cognitive/emotive abilities and tasks involved in moral judgment. What remains controversial, however, is the relevance of such neuroscientific insights as regards the conundrums of moral philosophy. Overstating the importance of neuroscience leads to naïve naturalism, by discounting the famous "is vs. ought" problem (a philosophical conundrum in its own right). On the other hand, declaring neuroscience altogether irrelevant is probably just as naïve [18]. Several authors note that neuroscience links up with moral philosophy in a rather indirect and problem-specific manner. J. Greene says that neuroscience provides a "behind the scenes" glance at how human morality functions, and believes that a better understanding of our moral instincts and their evolutionary origins may shed some light on philosophical controversies such as the clash between moral realism and moral subjectivism [18]. A. Roskies provides a thorough analysis of how results from neuroscience could be brought to bear on another long-standing philosophical dispute about the truth or falsity of "moral internalism" – roughly speaking, the assertion that moral judgments have a "built-in" motivational component which makes them fundamentally different from factual judgments. As a philosophical theory, internalism makes concrete assertions about what constitutes the essential nature of moral judgments, and some of these assertions have the potential to be confirmed or falsified by neuroscientific data [19]. In this particular case, relevant data may concern the extent to which the emotional and motivational component can be dissociated from the cognitive element of moral judgments.

Another classical area of philosophical controversy is the question of free will and responsibility. The development of neuroscience points to a purely causal account of brain events. One perceives, thinks, understands, and acts due to neural causes that have other neural causes themselves, or causes that lie in the outside world. The notion of a free

subject that initiates such causal chains seems unthinkable. This is at odds with the traditional "libertarian" and heavily metaphysical conception of freedom of the will. In this view, a free agent is supposed to act on *reasons*, not mere material causes. Furthermore, when an agent is faced with a choice between two courses of action and she chooses one, it can still be said that she could have chosen the other if she had wished (principle of alternative possibilities). Does neuroscience question this conventional wisdom, in turn challenging commonsense intuitions about responsibility and guilt? Does it force the judicial system to rethink its conceptual basis? This is a highly controversial topic today (the classical introduction to this problem is the 2004 article by J. Greene and J. Cohen [20]), but it seems that legal scholars and philosophers who believe that neuroscience is utterly irrelevant, just as those (mostly neuroscientists and other philosophers) who believe that neuroscience will bring a total conceptual revolution in legal and moral thinking, are probably both wrong.

The dilemma just discussed is often described as the free will vs. determinism problem, and the conventional mapping of possible philosophical opinions is defined by two questions: (1) Does free will exist? (2) Is determinism true? The most prevalent answers are "yes, no" (traditional libertarians, who believe that if determinism is true, genuine free will is impossible) and "yes, yes" (compatibilists, who think that some sort of free will can coexist with determinism). Interestingly, this does not exhaust all possibilities, because the question of whether there is free will could conceivably be independent of the question of determinism. This is the view defended by the British philosopher Galen Strawson. Very briefly, for Strawson, the impossibility of free will is basically a matter of logic:

"(1) When you act, you do what you do, in the situation in which you find yourself, because of the way you are.

It seems to follow that

(2) To be truly or ultimately morally responsible for what you *do*, you must be truly or ultimately responsible for the way you *are*, at least in certain crucial mental respects. (...) But

(3) You cannot be ultimately responsible for the way you are in any respect at all, so you cannot be ultimately morally responsible for what you do.

Why is it that you cannot be ultimately responsible for the way you are? Because

(4) To be ultimately responsible for the way you are, you would have to have intentionally brought it about that you are the way you are, in a way that is impossible." [21]

One seemingly obvious objection is that people can engage in "personal development" to change important aspects of their habits, personality, character, etc. But the reply is equally obvious: the fact that someone is or is not able to successfully change her mental nature is a reflection of that person's antecedent mental nature. In a sense, belief in free will is worse than false, it is unintelligible because it traps us into an infinite regress about what makes us be what we are and do what we do. Note that this argument is completely a priori and does not depend on some feature of the world such as the truth or falsity of determinism. In fact, indeter-

minism (for instance of the kind that seems implied by quantum physics) is equally damaging to the belief in free will as determinism, since an action that occurs for random reasons is no more free than an action that is necessitated by a deterministic string of causes and effects.

The point raised here does not attempt to discuss the philosophical dilemma of free will at any length. Rather, it simply tries to show that those holding different opinions on the problem of free will may well have different expectations – or a lack thereof – towards neuroscience. For instance, scientists with a philosophical agenda committed to a strong version of free will have tried to look for signs of “non-materialistic” mental causes [22]. Others, like Strawson, consider with good reason that their argument against free will is a genuinely a priori argument. Logically, they may not expect neuroscience to add much to it, except perhaps to uncover facts that bolster the intuitive force of the argument. The dialogue between neuroscientists, philosophers and ethicists is likely to be both fascinating and highly complex, beset with misunderstandings and differing expectations from both sides.

Back to the patient

The preceding discussion of neuroethics may seem quite abstract and far removed from the daily problems of medicine and health care confronted by neurologists, psychiatrists and forensic experts who deal with criminal responsibility. However, the fact that neuroethics shows promise as a distinctive and vibrant subfield of bioethics does not originate exclusively with its special relationship with philosophy. It also reflects the global importance of health problems involving the brain-mind. Whether they are defined as neurological or psychiatric diseases, the global burden of these pathologies is growing, underestimated, and massively undertreated, especially – but not only – in less affluent parts of the world. This was already underscored by the Global Burden of Disease study undertaken by WHO with the Harvard School of Public Health and other groups in the late nineties [23]. Furthermore, psychiatric and neurological diseases raise many more or less specific issues of social ethics, such as the fight against stigmatisation, the social integration of disabled persons in the workplace, fighting discriminatory practices in disability insurance and pension plans, or defending the place of diseases of the mind-brain in priority settings in health care. On a global scale, the problems of mental health are simply daunting. Recurrent episodes of catastrophic violence leave profound mental scars, which will need care long after the physical injuries have been healed, often in social contexts where psychiatric care is extremely scarce and has to be supplemented by lay counsellors (the example of Rwanda has been extensively studied, see [24–25]). Even so-called natural disasters, aggravated by poverty, lack of public infrastructures, and endemic violence, such as the floods in Pakistan today, will add mightily to the burden of mental illness for years and decades to come. If one adds the considerable problems of addiction as a medical, public health and judicial concern, it should be clear that neuroethics

needs a clinical and public health dimension, as well as a more theoretical and fundamental one. If neuroethics succeeds in providing a voice for these practical concerns within the neuroethics community, it will achieve what has been a constant aspiration of bioethics in general, namely to create a truly inclusive and interdisciplinary community of scholars and practical experts on topics of major philosophical, ethical and political importance.

Acknowledgement

I thank Prof. Samia Hurst for her insightful critical reading of the manuscript.

References

- 1 Jonsen AR. The birth of bioethics. New York: Oxford University Press; 1998.
- 2 Toulmin S. How medicine saved the life of ethics. *Perspectives in Biology and Medicine*. 1973;25:736–50.
- 3 Lewens T. What is genetics? *J Med Ethics*. 2004;30(3):326–8.
- 4 Illes J, Kirschen MP, Edwards E, Stanford LR, Bandettini P, Cho MK, et al. Ethics. Incidental findings in brain imaging research. *Science*. 2006;311(5762):783–4.
- 5 Illes J, Borgelt E. Brain imaging: incidental findings: in practice and in person. *Nat Rev Neurol*. 2009;5(12):643–4.
- 6 Morris JC, Roe CM, Xiong C, Fagan AM, Goate AM, Holtzman DM, et al. APOE predicts amyloid-beta but not tau Alzheimer pathology in cognitively normal aging. *Ann Neurol*. 2010;67(1):122–31.
- 7 Morris JC, Roe CM, Grant EA, Head D, Storandt M, Goate AM, et al. Pittsburgh compound B imaging and prediction of progression from cognitive normality to symptomatic Alzheimer disease. *Arch Neurol*. 2009;66(12):1469–75.
- 8 Greely H, Sahakian B, Harris J, Kessler RC, Gazzaniga M, Campbell P, et al. Towards responsible use of cognitive-enhancing drugs by the healthy. *Nature*. 2008;456(7223):702–5.
- 9 Dennett DC. *Consciousness explained*. 1st ed. Boston: Little, Brown and Co.; 1991.
- 10 Kay KN, Naselaris T, Prenger RJ, Gallant JL. Identifying natural images from human brain activity. *Nature*. 2008;452(7185):352–5.
- 11 Naselaris T, Kay KN, Nishimoto S, Gallant JL. Encoding and decoding in fMRI. *Neuroimage*. 2010 Aug 4.
- 12 Haller S, Klarhoefer M, Schwarzbach J, Radue EW, Indefrey P. Spatial and temporal analysis of fMRI data on word and sentence reading. *Eur J Neurosci*. 2007;26(7):2074–84.
- 13 Seitz RJ, Franz M, Azari NP. Value judgments and self-control of action: the role of the medial frontal cortex. *Brain Res Rev*. 2009;60(2):368–78.
- 14 Nomi JS, Scherfeld D, Friederichs S, Schafer R, Franz M, Wittsack HJ, et al. On the neural networks of empathy: A principal component analysis of an fMRI study. *Behav Brain Funct*. 2008;4:41.
- 15 Roskies A. Neuroethics for the new millennium. *Neuron*. 2002;35(1):21–3.
- 16 Greene J, Haidt J. How (and where) does moral judgment work? *Trends Cogn Sci*. 2002;6(12):517–23.
- 17 Damasio H, Grabowski T, Frank R, Galaburda AM, Damasio AR. The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*. 1994;264(5162):1102–5.
- 18 Greene J. From neural “is” to moral “ought”: what are the moral implications of neuroscientific moral psychology? *Nat Rev Neurosci*. 2003;4(10):846–9.
- 19 Illes J. *Neuroethics: defining the issues in theory, practice, and policy*. Oxford: New York: Oxford University Press; 2006.
- 20 Greene J, Cohen J. For the law, neuroscience changes nothing and everything. *Philos Trans R Soc Lond B Biol Sci*. 2004;359(1451):1775–85.
- 21 Strawson G. *Free will*. In: Craig E, editor. *Routledge Encyclopedia of Philosophy*. London: Routledge; 2004.
- 22 Popper KR, Eccles JC. *The self and its brain*. New York: Springer International; 1977.
- 23 Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ. Global burden of depressive disorders in the year 2000. *Br J Psychiatry*. 2004;184:386–92.
- 24 Neugebauer R, Fisher PW, Turner JB, Yamabe S, Sarsfield JA, Stehling-Ariza T. Post-traumatic stress reactions among Rwandan children and adolescents in the early aftermath of genocide. *Int J Epidemiol*. 2009;38(4):1033–45.
- 25 Neuner F, Onyut PL, Ertl V, Odenwald M, Schauer E, Elbert T. Treatment of posttraumatic stress disorder by trained lay counselors in an African refugee settlement: a randomized controlled trial. *J Consult Clin Psychol*. 2008;76(4):686–94.