From: López-Silva, P., Valera, L. (eds) Protecting the Mind. Ethics of Science and Technology Assessment, vol 49. Springer, Cham. https://doi.org/10.1007/978-3-030-94032-4_4, pp.31-40.

Ethics and Neuroscience: Protecting Consciousness

Arran Gare

Introduction: Defining the Problem

The Hippocratic Oath is a code of ethics defining correct behaviour by physicians they are required to commit themselves to before being accepted into the profession. It was the first code of ethics for any profession. While originating in Ancient Greece, it subsequently evolved, but the current code still embodies many of the core injunctions of the original code. The most widely accepted current form is the 2006 *The Declaration of Geneva* by the World Medical Association to be taken before being admitted as a member of the medical profession. The most important of its injunctions are: 'The health of my patient will be my first consideration' and 'I will maintain the utmost respect for human life'. The first is a rewording of the injunction from Epidemics, Book I, of the Hippocratic school: 'Practice two things in your dealings with disease: either help or do not harm the patient.' This was later simplified to the most basic precept of the Hippocratic Oath: 'First, do no harm.' It is one of the principal precepts of bioethics that all students in healthcare that, given an existing problem, it may be better not to do something, or even to do nothing, than to risk causing more harm than good.

The development of neurotechnology could be subsumed with little modification under the Geneva formulation of the Hippocratic Oath, extending this precept to a commitment not to damage people's psychological health. On this precept, is should be very clear that the old practice of lobotomizing supposedly mentally ill patients, severing connections in the brain's prefrontal cortex and leaving them emotionally shallow, lethargic and unable to concentrate or take initiative, making it easier to manage chronically agitated, delusional, self-destructive, or violent patients, should have been ruled out when it was being practiced. Nowadays, there are far more interventions in the functioning of the brain available, and it is less clear what damage to psychological health would mean. Furthermore, with the development of neurotechnology, interventions in the future could go well beyond treatment of patients with neurological disorders. They could be used to 'improve' ordinary people. This makes it all the more important to characterize psychological health.

The Challenge of Mainstream Science to Ethics

The biggest problem is that as far as modern science is concerned, mind and consciousness are problematic concepts, while 'common sense' views are often vague and contradictory. This makes the notion of psychological health problematic. Many sciences, with the support of a good many philosophers, are committed to explaining away the mind and consciousness, or only allowing that consciousness is an epiphenomenon. A whole tradition of philosophy, originating with Thomas Hobbes, has striven to understand humans as nothing but complex machines. This has come to be identified with the scientific view of humans and has had a major influence on psychology. With the development of neo-Darwinism, molecular biology and information science, humans, as with other forms of life, have been characterized as machines for reproducing genes, where genes are understood as strings of DNA encoding information. The brain is then seen as an information processor, that is, essentially a computer. Humans can then be characterized as information processing cyborgs, with the brain being nothing but a carbon based computer. This is the conception of humans now being promoted by transhumanists, who argue that the extensions of humans through technology should be welcomed as an extension of what we are as humans, and in a more extreme

form by the posthumanists who argue that the whole idea of the human was a temporary aberration and should be abandoned.

From this perspective, if there is a place for health it would amount to not hindering the efficiency with which human organisms are able to process information and act efficiently on the basis of this information, and if possible, augmenting this efficiency. If parts of the body, including the brain, are seen as defective, there should be no problem with replacing them with artificial parts. Just as it is possible to provide amputees with artificial limbs, or people with defective hearts with artificial hearts, if the brain is defective in some way and cannot be repaired, or was defective to begin with, it should be possible to replace part of it with prosthetic parts. Some proponents of this view of life argue that in future it will be possible to download minds onto computers. If this is the case, it might be possible to replace the whole of people's brains with prosthetic brains, not only repairing defects, but greatly augmenting their power to retain and process information. Ordinary people will be able to far surpass the greatest chess masters of the present, and will be free of emotions which at present interfere with their efficiency.

If these arguments are correct, then with this conception of humans there should be no problem with traditional concerns about modifying the brain, such as concern with the effects of lobotomising patients to address their mental disorders, electro-convulsive and insulin shock therapy to cure their depression by destroying their memories, cutting the corpus collosum to cure epilepsy, or modifying people's moods with chemicals so they will be content with their current life. With the conception of the brain promoted by information scientists, neurotechnologists are entirely justified in attempting to modify people's brains, possibly by removing bits and adding artificial components in order to make them conform to social conventions and think and act more efficiently. In fact, such procedures could be defended on the grounds that this will make humans more competitive with the robots that will be manufactured incorporating new advances in artificial intelligence.

The lesson that should be learnt from this is that the code of ethics that should be adopted in neurotechnology depends almost entirely on how the mind and the brain and their relationship are understood. At present, it is reductionist science culminating in the mechanization of the mind by cybernetics and information science that are taken to be the cutting edge of science and are being embraced not only by scientists but also by philosophers.

However, this raises another issue. Is this triumph of cybernetics and information science due to their having proved themselves to be the most promising research program, or because science itself is being corrupted. Funding comes from governments and increasingly, big business, who overwhelmingly fund the kind of science that will facilitate increased control over nature and people to advance military technology and/or generate more profits for corporations. The implicit goal is to replace humans as much as possible to reduce war casualties and labour costs, and to control or eliminate people who no longer have a place in this brave new world. There are now a number of works showing this to be the case, with governments forcing academics to obtain their funding from business corporations to ensure that it is only this kind of science that is funded. If this is the case, what is required is not only a code of ethics for neurotechnology, but a code of ethics for science itself to prevent its corruption.

But then the problem could be not just the corruption of science, but with science as such. The commitment to explanation involves a commitment to reductionism, since explanations imply showing that appearances are nothing but the effects of something else. Following this logic, the ultimate explanations will be in terms of the basic existents of the universe. These used to be thought of as elementary particles or force fields, but information has now been added to these. This trajectory

and its consequences were foreseen by Martin Heidegger. As he wrote in 'The Question Concerning Technology' (1977, 21):

> Modern science's way of representing pursues and entraps nature as a calculable coherence of forces. Modern physics is not experimental because it applies apparatus to the questioning of nature. The reverse is true. Because physics, indeed already as pure theory, sets nature up to exhibit itself as a coherence of forces calculable in advance, it orders its experiments precisely for the purpose of asking whether and how nature reports itself when set up in this way.

While initially, the subject was privileged as a non-objective being in control of science, the ends up being dissolved by objective science. As Heidegger wrote in 'The Age of the World Picture' (1977, 152f.):

> In the planetary imperialism of technologically organized man, the subjectivism of man attains its acme, from which point it will descend to the level of organized uniformity and there firmly establish itself. This uniformity becomes the surest instrument of total, i.e., technological, rule over the earth. The modern freedom of subjectivity vanishes totally in the objectivity commensurate with it.

The rise of cybernetics and the triumph of information science committed to total control of the world is the inevitable outcome. As Heidegger observed in 'The End of Philosophy and the Task of Thinking' (1978, 375f.):

> No prophecy is necessary to recognize that the sciences now establishing themselves will soon be determined and steered by the new fundamental science which is called cybernetics. ... For it is the theory of the steering of the possible planning and arrangement of human labour. Cybernetics transforms language into an exchange of news. The arts become regulated-regulating instruments of information.

So long as we accept this conception of science, the idea of a code of ethics for anything, let alone a code for neurotechnology, is problematic. If a code of ethics is to be defended for anything at all, it is necessary to re-open the question, What are humans? and What is science?

Philosophical Anthropology, the Humanities and Post-reductionist Science

These questions cannot be answered from within science by itself. They can only be answered with reference to the humanities. Since Plato, the question what are humans has always been at the centre of philosophy and the basis of the humanities. It was central to Aristotle's philosophy and it was central to Hobbes' philosophy in his effort to replace Aristotle's conception of humans. While Hobbes' philosophy was entrenched in culture through the scientism that he had defended, implying that only mechanistic science produces genuine knowledge, his work problematized the subject and subjective experience. While empiricists, granting a place to sense experience, attempted to uphold scientism, their efforts to do so were undermined by their assumptions. In the last paragraph of his book *An Inquiry into Human Understanding,* David Hume (1955, 173) concluded:

> When we run over libraries, persuaded by these principles, what havoc must we make? … [L]et us ask, Does it contain any abstract reasoning concerning quantity or number? No. Does it contain any experimental reasoning concerning matter of

> fact and existence? No. Commit it then to the flames, for it can contain nothing but sophistry and illusion. This injunction would require the reader to cast his own book into the flames.

It was in response to such work that Kant was inspired to make philosophical anthropology the focus of his philosophy. In his *Introduction to Logic* (2005, 17), published in its final form in 1800 and which guided his critical philosophy, Kant proclaimed that philosophy in its cosmic sense 'is the only science which has a systematic connection, and gives systematic unity to all the other sciences.' It can be reduced to four questions, What can I know? What ought I to do? What may I hope? and What is Man?, and Kant concluded 'all these might be reckoned under anthropology, since the first three questions refer to the last.'

It should be noted that while philosophical anthropology is made central to philosophy, it is inseparable from other domains of philosophy. If the question What can I know? can only be answered with reference to philosophical anthropology, the claim of philosophical anthropology to supply knowledge presupposes an answer to the question What can I know? Similarly, to engage in efforts to achieve such knowledge in order to work out What ought I to do? already presupposes that we know what we ought to do – engage in such efforts. The focus on philosophical anthropology made these interconnections clear, and appreciation of this was central to all Kantian, neo-Kantian and post-Kantian philosophy, including hermeneutic philosophy and phenomenology. What Kant showed was that the conception of humans put forward by Hobbes and the empiricists was too impoverished to account for the possibility of science. To account for science, we have to recognize the creative role of the subject in perception and in acquiring knowledge, requiring much more robust notions of imagination, reasoning and agency than the mechanistic view and the empiricism it engendered could countenance. It is also necessary to accord a place to the human capacity for autonomy, without which, all apparent beliefs would have to be viewed as epiphenomena of physical processes and no better or worse than any other beliefs, except in so far as they provide an advantage in the struggle for survival by what are now characterized as 'gene machines', machines for reproducing DNA.

Philosophical anthropology has been the thread running through what analytic philosophers deride as 'continental philosophy', having been developed by Kant's students, Herder and Fichte, and then by the Early Romantics and Idealists such as Hegel, then through to hermeneuticists such as Dilthey, neo-Kantians such as Ernst Cassirer, by the pragmatists, and many of the phenomenologists. These neo-Kantian and post-Kantian philosophers emphasised the essential social nature of human consciousness, that humans only develop the capacity for freedom through viewing themselves from the perspective of others and through being formed by their cultures, and generally, they promoted an ethics based on the notion of mutual recognition of each other's freedom, self-realization as participants within communities, and recognition of the intrinsic value of life.

Friedrich Schelling was exemplary in this regard, arguing that humans conceived as such have to be understood as having evolved within nature. If there is a clash between this conception of humans and Newtonian physics, then physics will have to be transformed. Accepting Kant's argument that we organize our experience through imagination and concepts, but rejecting Kant's claim that through transcendental deductions it canbe shown that we have to accept the concepts of prevailing physics, he argued that we can criticise and replace defective concepts and thereby bring nature, and humanity as part of nature, to a higher state of consciousness of itself through us. To this end, he argued for a philosophical physics in which activity, later characterized as energy, is fundamental, and characterized matter in terms of forces, arguing that this new physics would make magnetism, electricity and light and the relationship between them intelligible. He also argued for the development of new mathematics adequate to this more dynamic view of nature. On the basis of

these concepts he argued for an evolutionary cosmology granting a place to emergence through the limiting of activity. Emergent entities might appear as objects, for instance crystals or chemicals of various kinds, but Schelling argued, these should be seen to be products of the activity of opposing forces achieving a balance. They are emergent, and to some extent immanent causes of themselves, and this makes it impossible to explain them as merely the effects of their environments and constituents. In chemistry, these opposing forces are now referred to as valances which generate molecules of various complexity and stability. Schelling characterized the distinctive characteristics of living beings as processes that must actively maintain their form while interacting with their environments, so these environments are defined in relation to them as their worlds. With this characterization of life, it was then possible to characterize and explain the distinctive characteristics of humans as essentially social, self-conscious beings living in culturally constituted worlds, capable of understanding their own history within the context of the history of nature and reflection on and transforming their cultures. In all cases, living beings, including humans, are inseparable from their environments, but are to some extent the immanent causes of themselves.

These are the ideas which triumphed with the development of thermodynamics and the field theories of electro-magnetism of Faraday and Maxwell, with the development of chemistry and then relativity theory showing that matter is really a form of energy. While mainstream biology is reductionist, reductionism has been shown to be incoherent (matter can't evolve) and is strongly challenged by holistic ideas associated with systems theory, including the theory of complex adaptive systems and anticipatory systems theory, process metaphysics, hierarchy theory, biosemiotics, and efforts to account for consciousness using quantum field theory (Vitiello, 2002; Ho, 2004, 228ff.). All these are part of the anti-reductionist tradition of thought and research program inspired by Schelling and those he influenced (Gare, 2013). These are the forms of thinking being advanced in modern science that are consistent with work in philosophical biology and philosophical anthropology. Advances in science have produced what Ilya Prigogine and Isabelle Stengers (1984, p.xxixf.) called 'the new alliance' between science and the humanities.

The Clash Between Reductionist and Anti-Reductionist Science

Most scientists do not accept this post-Newtonian research program, however. They still promote particle physics or string theory rather than accept the advances in quantum field theories, and claim that statistical mechanics as developed by Ludwig Bolzmann has displaced thermodynamics. And they fail to appreciate the modern chemistry and nuclear physics are triumphs of Schelling's post-reductionist thinking. On the basis of their acceptance of statistical mechanics, ignoring its limitations (for instance, accounting for phase transitions, let alone the dissipative structures examined by Prigogine) they embrace Boltzmann's notion of entropy and equate negative entropy with the notion of information developed by Shannon to analyse the capacity of cables to transmit messages. As noted, combined with cybernetics, this notion of information provides the foundation of information science, and it is claimed to be able to account, along with molecular biology, for life and mind.

These views are not accepted by advanced theoretical biology, however. Jesper Hoffmeyer in *Signs of Meaning in the Universe,* essentially a manifesto for biosemiotics based on Peirce's philosophy, pointed out that 'form' for the Romans was a mangled version of the Greek 'morf' (or 'morph'), and 'information' meant being formed mentally. Atomistic thinking in the Twentieth Century led 'information' to be understood as isolated chunks of knowledge and this was taken over by the physicists, who then characterized it as something in the world, independent of anyone, and then tried to impose this inverted, desiccated concept of information on all other disciplines. In his later book *Biosemiotics,* he wrote that 'up-to-date biology must acknowledge that the biochemical concept

of information is just too impoverished to be of any explanatory use' (p.61). As far as the computational notion of the mind is concerned, as Jeremy Fodor (2000) pointed out, the mind does not work that way.

Is it just a matter of choice between rival research traditions? My contention is that it is not. The tradition inspired by Schelling is far more coherent and has proved far more fruitful than the rival reductionist tradition, even when this reductionist tradition utilizes concepts such as fields (in bowdlerized form) inspired by the Schellingian tradition and incorporates the notion of information. The post-Newtonian tradition can make intelligible whatever advances have been made through reductionist approaches in the sciences, while reductionist approaches cannot make sense of what is comprehensible from the anti-reductionist research tradition, including the existence of ourselves as conscious beings. Reductionism is the dominant tradition because science has been corrupted. Firstly, by those who fund science who are for the most part only interested in knowledge that facilitates control over nature and people. This is what reductionist science delivers. It is based on controlling situations and modifying components to enable predictions to be made, that is, as Heidegger observed, enframing the world to reveal it as standing reserve to be controlled and exploited. Secondly, it is far easier to develop such science. Following the 'scientific method', ultra-specialists add small increments to the bucket of scientific knowledge. This can be real knowledge, but trivial. Through specialization, these scientists lose contact with other branches of science and with an integral interpretation of the universe, and this is really the negation of science.

It has been shown that scientists who have done most to advance science have been characterized by a wide range of interests which they take very seriously, including interest in the arts and humanities (Root-Bernstein, 2015). It is for this reason that they can go beyond established methods and ways of thinking and develop new concepts and create new methods. Both politicians and ultra-specialists are hostile to such scientists who are also prone to speak out on matters of public interest and who, by questioning the assumptions of mainstream science, challenge the work and career prospects of these ultra-specialists. When politicians, business leaders, managers of universities and research institutions ally themselves with such ultra-specialists and attempt to manage science, demanding quantifiable outputs, science stagnates. It has been strongly argued by a prominent medical researcher, Bruce Charlton (2012) that this is the case with current science. Furthermore, Joseph Ben David (1971) showed that throughout history from the Greek onwards, whenever governments have tried to control science to channel it to serve their interests, even when they have increased support for science, they have destroyed its creativity. For real science to flourish, there has to be an 'autonomization' of the 'scientific field', as this was characterized by Pierre Bourdieu (2004), so that truth versus falsehood becomes the basis for competition and having the conditions to advance knowledge, rather than patronage, usefulness and ability to get research grants. This problem is particularly acute with neuroscience where, given the current state of science and its severance from natural philosophy, those who are likely to gain research support are the hyper-specialists aligned with reductionist science, unwilling to consider the problematic relationship between their scientific research and the reality of conscious experience, and willing to serve whoever pays them.

It is because the scientific field, and more generally, the academic field have been corrupted, and in this corrupt state are imposing a nihilistic world-view that totally devalues life, that developing a code of ethics for neurotechnology important, and problematic. It is first necessary to have a code of ethics for science. What is needed of such a code is for all participants in the scientific endeavour to uphold the autonomy of the scientific field as the condition for the flourishing of science. Above all, this involves upholding the quest for truth and the conditions for those who are engaged in this quest, and to sustain this, reflexivity on the part of scientists about their own enterprise and the conditions

for it. The quest for truth should be understood as the quest for a comprehensive understanding of the cosmos, of life and humanity, and the place of humanity, including science and scientists, in the cosmos, and all specialist inquiry should be related to this quest. What is important is that those accredited as scientists have a deep commitment to advancing our understanding of the world, which requires of them knowledge of the history of both science and natural philosophy. This involves rejecting the idea that science can be treated as a mere instrument accumulating useful knowledge, and acknowledging that the health of science requires recognition that the scientific field has its own immanent dynamics that must be respected, giving autonomy to scientists who, by virtue of their drive to comprehend the world, are unpredictable. The value of their work cannot be quantified and managed on the basis quantifiable indicators. That is, the conditions for autonomous enquiry must be respected and cultivated.

What is central to creating a code of ethics for science, is central to all ethics and all professions, and it pinpoints the central problem that has to be overcome. The corruption of science occurs through enframing the scientific community and its individual members as standing reserves to efficiently exploited, reducing them to nothing but instruments for advancing useful technology. It is through this enframing that those managing science have kept in ascendancy utterly debased notions of humans as stimulus-response mechanisms or information processing cyborgs in place of more philosophically and scientifically defensible notions of humanity. With such enframing, efficiency is the only evaluative criterion to judge science and scientists, and ultimately, even this is undermined. The development of neurotechnology brings this dilemma into sharp focus. Human brains are being treated as standing reserves to be efficiently controlled and exploited. However, the advance of science requires that the autonomy of the scientific field and scientists be respected so that the field can develop according to its own immanent dynamics. Since the telos inspiring these immanent dynamics is a coherent understanding of the world, including the place of science and scientists in this world, this will involve defending a conception of humans that acknowledges that they also have autonomous dynamics not completely explicable in terms of their environments and constituents. Autonomous science should itself provide the basis in the conception of humanity developed and defended for valuing and defending the autonomy of scientists.

What is required above all of a code of ethics is recognition that individuals are autonomous agents. As Immanuel Kant put it (1959, 47): ‘Act so that you treat humanity, whether in your own person or in that of another, always as an end and never as a means only’, where being an end was associated with having the capacity for autonomy. Accepting this principle, the only acceptable intervention in either the biological or social conditions of people is that it does not damage their autonomy, but if anything fosters it. And this is especially the case with neurotechnology. Going back to the Hippocratic Oath and how this should be extended to deal with neurotechnology, the injunction ‘I will maintain the utmost respect for human life’ can be taken to imply respect for people’s autonomy. ‘The health of my patient will be my first consideration’ can be taken to imply ‘The capacity for autonomy of my patient will be my first consideration.’ Summarizing this to conform to the traditional formulation, ‘First, do no harm’ can be reformulated as ‘First, do not undermine the capacity of people for autonomy.’ And what is right in a medical context is right everywhere. Neurotechnology should never be deployed to undermine people’s autonomy, and should only be deployed to augment people’s autonomy.

References

Bourdieu, Pierre. 2004. *Science of Science and Reflexivity*. Chicago: University of Chicago Press.

Charlton, Bruce G. 2012. *Not Even Trying: The Corruption of Real Science*. Buckingham: University of Buckingham Press.

Fodor, Jerry A. 2000. *The Mind Doesn't Work That Way*. Cambridge, Mass.: MIT Press.

Heidegger, Martin. 1977. 'The Age of the World Picture'. In: *The Question Concerning Technology and Other Essays*. Trans. William Lovitt. New York: Harper Torchbooks.

Heidegger, Martin. 1978. 'Letter on Humanism'. In: *Martin Heidegger: Basic Writings*, Ed. David Farrell Krell. Trans. Frank A. Capuzzi. London: Routledge & Kegan Paul.

Ho, Mae-Wan. 2004. *The Rainbow and the Worm: The Physics of Organisms*. 3rd ed. Singapore: World Scientific.

Hume, David. 1953. *An Inquiry Concerning Human Understanding*. Indianapolis: Bobbs-Merrill.

Kant, Immanuel. 1959. *Foundations of the Metaphysics of Morals*. Trans. Lewis White Beck. Indianapolis: Bobbs-Merrill.

Prigogine, Ilya & Stengers, Isabelle. 1984. *Order Out of Chaos: Man's New Dialogue With Nature*, Toronto: Bantam Books.

Root-Bernstein, Robert. 2015. 'Arts and crafts as adjuncts to STEM education to foster creativity in gifted and talented students,' *Asia Pacific Education Review*, 16(2): 203-212.

Vitiello, Guiseppe. 2002. *My Double Unveiled: The Dissipative Quantum Model*. Amsterdam: John Benjamins.