Towards Neurobioethics

Towards Neurobioethics

By

Darlei Dall'Agnol

Cambridge
Scholars
Publishing

Towards Neurobioethics

By Darlei Dall'Agnol

This book first published 2020

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-5801-0
ISBN (13): 978-1-5275-5801-4

To Fred, Peter and Steve

CONTENTS

ABBREVIATIONS

ADD	Attention Deficit Disorder
ADHD	Attention Deficit-HyperactivityDisorder
ANVISA	Agência Nacional de Vigilância Sanitária (BR)
APA	American Psychiatric Association
BCIs	Brain-Computer Interfaces
B.R.A.I.N. Initiative	Brain Research through Advancing Innovative Neurotechnologies
CBA	Cost-Benefits Analysis
CEA	Cost-effectiveness Analysis
CNS	Central Nervous System
DBS	Deep Brain Stimulation
DIY tDCS	(Do It Yourself) Transcranial Direct Current Stimulation
DLPFC	Dorsolateral Prefrontal Cortex
DSM-V	Diagnostic and Statistical Manual of Mental Disorders
ICD	International Classification of Diseases
IRBs	Institutional Review Boards
FDA	Food and Drug Administration
FEP	First Ethical Principle
fMRI	Functional Magnetic Resonance Imaging
MR	Magnetic Rosonance
O3	Our Brain, Our Selves, Our World: An International Research Consortium on Neuroscience, Ethics and Society
PCE	Pharmaceutical Cognitive Enhancers (PCE)
PET	Positron Emission Tomography
PFC	Prefrontal Cortex
RTW	Rational Theory of Welfare
tDCS	Transcranial Direct Current Stimulation
TMS	Transcranial Magnetic Stimulation
UDHR	Universal Declaration of Human Rights
VMPFC	Ventromedial Prefontal Cortex
WM	Working Memory

Some Philosophical Works (see references):

NE	*Nicomachean Ethics* (Aristotle)
GMM	*Groundwork of the Metaphysics of Morals* (Kant)
TLP	*Tractatus Logico-philosophicus* (Wittgenstein)
PE	*Principia Ethica* (Moore)
PI	*Philosophical Investigations* (Wittgenstein)
UTIL	*Utilitarianism* (Mill)

ACKNOWLEDGMENTS

In the first semester of the academic year 2017-8, I presented a three-part lecture series entitled "Respectful Care in Bioethics and Neuroethics," co-sponsored by the Department of Philosophy and the Center for Ethics and Humanities in the Life Sciences, at the Michigan State University in East Lansing, US. For the lectures, I used my paper "Knowing-how to care" and my book "Care and Respect in Bioethics" (see References) as the main bibliographical sources. At this time, I also did some research on the foundations of neuroethics to investigate the practical implications of the concept of *caring respect*. This book is a revised version of those lectures, which I delivered as a Visiting Scholar supported by the Fulbright Foundation. Thus, I am grateful to Fulbright for funding this challenging journey and for their work to improve the academic relationship between Brazil and the US.

I would also like to thank the participants in the lectures, especially David Godden, Fred Gifford, Todd Hedrick, Matt Ferkany, Laura Cabrera, Robyn Bluhm, Jamie Nelson, Matthew W. McKeon and Professor Peter Railton (who did me the honor of coming from Ann Arbor), and the students for their questions, comments, suggestions etc. and for our conversations over lunch and on many other occasions. Thanks also to Nicole Carolyn-Indira Shilling for inviting me to present my ideas at the MSU Philosophy Club. Many thanks especially to Professor Frederick Rauscher for organizing the lectures at MSU and for our ongoing philosophical discussions, not only on Kant's meta-ethics and the Kantian bioethical principle of respect for persons, but on some Rawlsian political issues too.

I would also like to thank some colleagues of the current research group I am working with, namely the *Global Consortium O3*: Our Brain, Our Selves, Our World (An International Research Consortium on Neuroscience, Ethics and Society.), coordinated by Professor Ilina Singh (Oxford University), especially those dealing with the neuroethics panels, namely Professors Julian Savulescu and Walter Sinnott-Armstrong for their comments to my "case-response" in our meeting at the Ashridge Business School (England, June 2016), supported by the Welcome Trust. I am also grateful for the opportunity to meet and to discuss some points with the North American neuroethicists, particularly Martha Farah and Steven Hyman. I was invited to be part of a thematic group in this consortium to contribute with my expertise in bioethics having done some research at The

Oxford Uehiro Centre for Practical Ethics. The ideas defended here of course do not represent the official view of the O3, but only my personal understanding of such matters.

INTRODUCTION

Reshaping the Field of Neuroethics

From time to time, a particular science achieves such great success that people are tempted to elevate it to the condition of *prima philosophia* and then to try to explain everything else from its perspective. Thus, physics becomes physicalism; history, historicism, and so on. Nowadays, the big science is the investigation of the nervous system, particularly *the brain*, as can be recognized by looking at the US B.R.A.I.N. Initiative (Brain Research through Advancing Innovative Neurotechnologies); the EU's Human Brain Project; Brazil's IINNT (a campus of the brain in Natal, RN), idealized by Miguel Nicolelis; the Institute of the Brain, which is also in Brazil (Porto Alegre, RS) and led by the Argentinian-born Ivan Izquierdo, etc. The main paradigm is, then, given by neuroscience and everything else seems to require its prefix: neurophilosophy, neuroeconomy, neuroeducation, neurotheology, neurolaw, neurotechnology, neuroethics, neuropolitics etc. I do not need to multiply examples to show that "neuromania" is upon us; that is, the belief that one needs to learn the workings of the brain to know anything at all. Thus, we seem to live in the age of "brainism" (Bakhurst, 2008). But what does it really mean to use "neuro" as a prefix to a word as it appears in the title of this book?

To answer this question and to clarify the main argument of this work, I would, in this introduction, like to present some of the general philosophical assumptions that guide my work both in bioethics and neuroethics and to propose reshaping the field of neuroethics towards what I here call "neurobioethics." My general argument can be put forward in this way:

1º–since the predominant use of "neuroethics" is ambiguous and may involve categorial and other errors in trying to extract normative implications directly from neuroscientific studies, the ethics of neuroscientific experiments must be based on classical bioethics despite the fact that we need also to overcome the intuitionist meta-ethical problems of the predominant normative theory using a new moral epistemology, namely *practical cognitivism*;

2º–now, despite the fact that classical bioethics is already well-armed with the main ethical guidelines to perform neuroscientific

experiments on human beings in the right way, especially by applying the *prima facie* principle of respect for persons, it also needs some normative improvements towards a more unified account of *care* and *respect* since this unification can lead, for instance, to the concept of *caring respect*, which is crucial for a common and shareable morality in the public sphere;

3º–thus, by applying *practical cognitivism* and *caring respect* to the ethics of neuroscientific investigations and their associated neurotechologies, for example to the moral problems of cognitive enhancement using nootropics, we can better recognize that neuroethics is just a subfied of bioethics and not an independent branch of applied ethics; hence, *neurobioethics*.

Each of these steps in the argument will be made in the following three chapters; the first will be on the *meta-ethical* basis of neurobioethics, the second on the *normative* foundations of neurobioethics, and the last on some *practical* issues belonging to neurobioethics. In the end, the entire argument will reach the conclusion that we need to depart from neuroethics, unduly compromised with scientificism and utilitarianism, *towards neurobioethics*.

To explain in greater detail what neurobioethics is all about, I would first like to present some methodological remarks. In general, I will follow a Wittgensteinian orientation as given by the neurologist M. R. Bennett and the philosopher P. M. S. Hacker, who, in their excellent book *Philosophical Foundations of Neuroscience*, made clear what philosophy and neuroscience can do and what they cannot do: the former proceeds through *conceptual analyses*; the latter by doing *empirical investigations* (2003, 378f). Conceptual analysis is, then, in some sense *prior* and not reducible to empirical investigations (the sciences). Moreover, I will argue that Bennett and Hacker are right in holding that science must avoid being committed to metaphysical reductionism (2003, 355f). I will expand on this point soon.

Now, one remark I would like to make straight away is that the *neuroscience* of morality (sometimes used as one of the meanings of "neuroethics") is a worthy kind of investigation, but it cannot, by itself, provide a normative framework for dealing with research using humans as subjects. To accomplish that, we need (bio)*ethics*, which is an autonomous area of study. To think otherwise is to fall into scientificism; that is, into a metaphysics that leads to a partial and distorted understanding of reality. As Wittgenstein puts it (TLP, 6.52), we feel that even when all possible scientific questions have been answered, the problems of life remain untouched. Thus, the general aim of this section is to offer a broader picture

of ethics as a *philosophical* enterprise in order to provide a background for the argument I will present and defend throughout this work.

I would, then, like to start with a clear definition of philosophy itself to make explicit what it can and cannot do, and to establish the foundations of neurobioethics. I take pure philosophy to be just *a search for wisdom*. In this sense, I would like to honor Potter's original bioethical project: "Mankind is urgently in need of new wisdom that will provide the 'knowledge of how to use knowledge' for man's survival and for improvement in the quality of life." (Potter 1971, 1) The author of the book *Bioethics, Bridge to the Future* goes on to clarify that the humanities and philosophy, understood as "love of wisdom," play an important role, together with the science of biology, in building the "science of survival" he calls "Bioethics." As can be seen, Potter emphasizes the two most important ingredients in achieving this new wisdom. Granted, philosophy is a search for wisdom, but what is *wisdom* after all?

I believe Potter was on the right track when he envisaged a kind of metaknowledge, namely *the knowledge of how* to use scientific knowledge for *ethical* purposes, rescuing in this way wisdom's *protagonist* role in our lives. That is to say, there are many different kinds of knowledge and philosophy cannot be conflated with any type: it is neither *common sense* nor *one of the natural or social sciences.* Philosophy is not common sense since it is not a-methodic and uncritical in the way common sense is, working through trial and error only; it is not full of unexamined beliefs or values; it is not "popular wisdom" either. On the contrary, philosophy is most critical of some conventional beliefs and practices.

Philosophy is not a particular natural science either since it does not proceed in experimental ways using empirical methods (induction, hypothetical-deductive, etc.) *to explain* causally natural phenomena, but rather works conceptually and argumentatively; it has no particular object, but everything is open to its reflection; it is *knowing-how to use scientific knowledge to improve the quality of our lives*. Nor is philosophy one of the social sciences since it does not attempt to understand or *to describe*, for instance, different moral traditions. Moreover, philosophy has, since classical times, been driven by ethical imperatives such as *Know thyself*! leading to the re-*cognition* of the limits of scientific knowledge itself. Thus, philosophy is non-naturalistic and its special status can be shown as follows: despite the fact that empirical sciences such as biology are the best way of explaining causally many facts concerning the world around us, science needs philosophy, both to refute radical skepticism –thus instituting the very possibility of scientific knowledge itself– and to establish ethical ends for its activities. The sciences cannot establish themselves in a circular way,

that is, by asserting scientifically that there is scientific knowledge. To achieve this aim, we need a different kind of understanding, namely a philosophical one which shows that skepticism is nonsensical. Besides, both science and neurotechnology can be used to kill or to save lives, so it is paramount to ground them on a sound philosophy to ensure, to use Potter's words, man's survival and an improvement in quality of life, which is an *ethico-philosophical* end, not a scientific one.

It is also crucial to distinguish philosophy from the arts and the religions, especially from ideologies and other forms of sophistry. The arts and religions are *expressions* of human imagination. Religions are also based on non-cognitive grounds such as *dogma* and *faith* and feelings (e.g., fear and hope); on the contrary, philosophy cares for wisdom. Most religions postulate supranatural entities such as God, souls etc., while a sound philosophy cannot commit itself to such an ontology. An attempt to pass off a specific religion as scientific is an ideological one, as well as being an effort elevate science to metaphysical significance. An ideology is a set of beliefs and values that obscures part of reality; it disguises wishes and interests. Philosophy is above all committed to truth and to emancipation from false or hidden beliefs. It is not any sophist self-proclaimed possession of the ultimate truth, but a permanent search to improve wisdom. Besides, only philosophy is based on universal human values, whereas ideology and sophistry are based on specific ones.

A philosophy based on wisdom has an irreducible ethical component, and if the previous remarks are cogent, then it is easy to grasp the *autonomy* of ethics; namely, it is neither a science nor a religion. To conflate values (ethics) and facts (sciences) is to commit the naturalistic fallacy by defining non-natural concepts such as "good," "right" etc. in natural terms (e.g., more evolved). This is a category mistake which thus invalidates all conclusions. I will return to this point while discussing the problem of crisscrossing language-games (chapter 1).

Now, to avoid similar confusions, it is crucial to make the distinction between *ethics* and *morality*. Basically, I differentiate "ethics" as a philosophical discipline from "morality" as referring to particular value systems. First, then, I will sort out the main domain of philosophical ethics, and I will then provide a better definition of "morality" soon. Afterwards, I will construct a map of applied ethics to locate neurobioethics.

Ethics is normally seen as comprising three main spheres:

(i) Meta-ethics;
(ii) Normative Ethics and
(iii) Applied Ethics.

I will deal with each of these domains in the three chapters of this book, which sort out respectively the three main ideas of this work: one is meta-ethical (practical cognitivism as a substitute for the intuitionist basis of neurobioethics); another is normative (the concept of caring respect) and, finally, there is a practical one (the permissibility or not of cognitive enhancers) which is closely related to the three steps of the main argument as stated above.

Meta-ethics is, then, an examination of the epistemic and ontological status of ethics itself. For instance, some meta-ethical issues are whether moral knowledge exists (cognitivism *vs.* non-cognitivism); whether there are moral proprieties and whether they are natural ones (realism *vs.* anti-realism); whether morality is intrinsically motivating and so on. In chapter 1, I will develop the main tenets of a new moral epistemology to justify some meta-ethical and normative claims I make later, for instance, that one *knows-how* to respect a vulnerable subject of neuroscientific research (or not). Thus, I will argue that there is moral knowledge and that it is best understood as involving knowing-how to act, which goes under the label of "practical cognitivism" and will provide the meta-ethical grounds for neurobioethics.

Now, *normative ethics* deals with ways of distinguishing right from wrong, good from bad etc. in the moral sense. Normative ethics is currently divided into three main normative approaches: (i) *virtue ethics* (right is what a virtuous person would do in the circumstances); (ii) *consequentialism* (right is what brings the best results) and (iii) *deontology* (right actions may have worth in themselves and be prescribed by universal norms). In chapter 2, I will try to show that there is a way of overcoming this tripartite division, making, for instance, an ethics of care and an ethics of universal respect compatible and congruent, thus leading to attitudes such as *caring respect* which may have important applications in all domains of life, including in neurobioethics.

Applied ethics is the practical domain where particular moral issues are analyzed using a normative approach, for instance, bioethical problems such as whether voluntary euthanasia is permissible; neuroethical problems such as whether cognitive enhancement by pharmaceutical means is morally acceptable for teenagers; roboethical issues such as whether artificial agents have rights; ecoethical issues such as whether biodiversity has intrinsic worth; zooethical issues such as whether non-human animals can be used in neuroscientific experiments; etc. In chapter 3, I will discuss neuroethical problems in great detail, especially cognitive enhancement. As can be seen, there are apparently many subfields within applied ethics: bioethics, neuroethics, roboethics, ecoethics, zooethics, etc., which deal

with practical, day-to-day moral problems using *normative* approaches grounded on *meta-ethics*. I will better clarify the interconnections between these subfields of applied ethics below, explaining why we need to walk in the direction of neurobioethics.

From a philosophical point of view, meta-ethical problems have priority. That is to say, a philosopher must first make sure that the meta-ethical bases are well settled in order to find the best normative account to address practical concerns. It is almost unnecessary to say that applied ethics has priority in our everyday lives. I will deal with some practical problems in this book, especially issues surrounding the use of drugs for cognitive enhancement as well as treatment to illustrate the meta-ethical and normative points I propose here.

Having made these distinctions of the main domains of ethics, it is now possible to better distinguish it from *morality*. Consider Roger Crisp's definition of a morality (2006, 9):

> … a set of cognitive and conative states, including beliefs, desires, and feelings, which leads its possessors among other things to (*a*) view certain actions as wrong (that is, morally forbidden), and hence to be avoided, (*b*) feel guilt and/or shame as a result of performing such actions, and (*c*) blame others who perform such actions.

Thus, a particular morality can be understood as a value system, that is, a set of beliefs, practices, norms, virtues, feelings, etc., which lead us to regard particular actions as right or wrong. In this sense, there seem to be many value systems: Christian morality; liberal morality; Buddhist morality; communist morality etc. Whether this plurality of value systems implies relativism is an open question, as it is whether there are sufficient *shared values* to postulate "*the* common morality". I will discuss this point in chapter 2.

Now, it seems clear that we have to accept the plurality of value systems, and not only as a matter of fact. If the definition of philosophy as a search for wisdom is sound, and if wisdom has to do with the re-cognition of the limits of knowledge, then we are bound to accept *agnosticism* in metaphysics and its pluralist implications as a matter of principle. That is to say, specific moral systems may be based on *antagonic* metaphysical assumptions. For instance, one moral system may contain religious beliefs, while others can be built on naturalist values. Moral systems express, then, different worldviews, but we can only sketch a *finite* ontology. This finite ontology, though, may have important ethical purposes anyway. Thus, we may take a *methodological* naturalist approach to moral phenomena, but we cannot assume *metaphysical* naturalism in a reductivistic or eliminationistic

manner. This philosophical point has important practical consequences for what the sciences, for example neuroscience, can do to help us to understand morality. As we will see in this work, reshaping neuroethics in the direction of neurobioethics prevents us from committing a naturalistic fallacy and avoids many other problems.

Even if some degree of value pluralism should be accepted, not only as a matter of fact, but also in principle, it does not follow that radical moral relativism is true or that moral skepticism should be subscribed as a consequence. Normative relativism cannot be conflated with descriptive relativism. It is a *category mistake* to identify normative claims with matters of fact; that is, to try to infer an *ought* (a prescriptive statement) directly from an *is* (a descriptive proposition). From the idea that philosophy is a search for wisdom, we may extract the project of building up *a* common and sharable morality (henceforth, *a* CS-Morality) to make value pluralism reasonable. To rephrase the American political philosopher John Rawls (1993, xvii), there is no "the fact of a plurality of reasonable but incompatible comprehensive doctrines," but rather *pluralism must be made reasonable* through *a* CS-Morality. Such *a* moral system is the *sine qua non* condition for the pacific co-existence of and cooperation among divergent value systems. Therefore, reasonable value pluralism is only possible under the assumption of *a* CS-Morality. Radical ethical relativism turns out, then, to be false.

A similar argument can be presented to reject moral skepticism. As we saw above, there are many kinds of knowledge, and the denial of knowing-how, which presupposes the acquired capacity to follow norms, is simply self-contradictory. Actually, we can show that the skeptic presupposes what he wants to deny. This argument opens the door to constructing a moral epistemology based on the idea that there is moral knowledge and that it is best understood as involving knowing-how. This is the main purpose of chapter 1: to present the meta-ethical basis of neurobioethics and open the door to establishing a normative ethical theory. To illustrate, a person may know-how *to respect* other person, who happen to be a vulnerable being, namely in a *caring* way (or not). Once again, this meta-ethical claim leads us to the normative idea of *a* CS-Morality based on, as we will see in greater detail in chapter 2, *mutual care* and *reciprocal respect*.

I would now like to examine some interconnections between the subfields of *applied ethics*, especially the links between bioethics and neuroethics, and to present the domain of neurobioethics. I started this introduction by identifying Potter's project of bioethics as a bridge between the empirical sciences and the humanities to achieve a *new wisdom* that leads to an improvement in quality of life. Unfortunately, this is not the

predominant use of the word "bioethics" nowadays. In fact, its main employment relates to the principles and values of biomedicine and other health sciences only. As can be seen clearly, this is not what the very same Potter years ago called "global bioethics" in his later book, which includes many environmental issues and is built on Leopold's legacy, namely on the project of an ethics of land. Too bad for us that this is also not the main use of "global bioethics," since most bioethicists use it to refer to international bioethical issues. Thus, I would like to distinguish three main uses of "bioethics", referring to: (i) holist bioethics; (ii) special bioethics and (iii) global bioethics. We should bear in mind, however, that a truly *philosophical bioethics*, guided by the conception of philosophy spelled out above, maintains interconnections among all these domains.

How can we distinguish holist from special bioethics? I think that one way of seeing the differences is to refer to Potter's original project but to rename it *holistic bioethics* in the sense that it includes a moral concern for the environment and non-human animals too; that is, it *cares for life as a whole*. This fits well with what Fritz Jahr, perhaps the first person to have used the term "bioethics", had in mind when he proposed the following principle: "Respect every living being on principle as an end in itself and treat it, if possible, as such." (Jahr 2012, 4) Of course, there is the issue of whether we should respect every single living being or, as Leopold held, we should rather revere the biotic community as a whole. Leopold's ethical principle was, in his famous *A Sand County Almanac*, sorted out in these terms: "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community. It is wrong when it tends otherwise." (1993, 16) The implications of these principles, namely Jahr's and Leopold's criteria to distinguish right from wrong actions, may lead to different practices. To illustrate: the former seems to lead necessarily to vegetarianism or even to veganism; the latter, not necessarily. Unfortunately, I cannot pursue this issue here, but I would like to emphasize that *holist bioethics* would then amount to a general attitude of respect for life as a whole including the environment and non-human animals.

Now, *special bioethics,* in contrast*,* may refer exclusively to the principles and values needed to carry out scientific experiments that use human beings as subjects in the health sciences. It relates mainly but not exclusively to biomedical experiments. In fact, the classical "bioethical" book *Principles of Biomedical Ethics* starts by identifying them in its very first sentence: "Biomedical ethics, or bioethics, …" (Beauchamp & Childress 2013, vii). This identification is, however, misleading since it seems clear that there are deep interconnections between human well-being and a healthy environment. Moreover, an anthropologist does not do

research *on* human beings, but *with* human subjects. Unfortunately, this is still the predominant use, which gives us a very limited understanding of bioethics in general. Now, sub-fields of special bioethics include *clinical* bioethics, *neonatal* bioethics, etc. and arguably, *neuro*-bioethics itself. As we will see, neurobioethics is a subfield of special bioethics. Since I will, for didactic reasons, focus on *special bioethics* throughout this book, I will return to this point later to complement this exposition.

In addition, there is *global bioethics*, which deals with bioethical issues within international relations between countries. This is, for instance, what is mainly established in domestic or regional bioethical treaties, for instance, the Oviedo Convention, which prescribes norms for the European Community on many bioethical issues. With this in mind, we Americans should, perhaps, have an Interamerican Bioethics Declaration as we have for Human Rights (San José, 1969). There is, of course, The UNESCO Universal Declaration on Bioethics and Human Rights, but still, it seems justifiable to consider what is specific for the Americas as a distinct region. Now, some bioethicists are skeptical about whether we can reach a global consensus on such questions (e.g., Engelhardt 2006). In this work, I will briefly compare the North American and the Brazilian ways of doing bioethics (and neuroethics) in order to make sense of the idea of *a* CS-Morality, but I will not, for reasons of space, discuss all the issues related to global bioethics.

To avoid confusion and misunderstandings, we must keep in mind that all these distinctions are useful to the extent that they help us to focus on particular neurobioethical issues, but that they should not make us lose sight of the forest. Thus, I have proposed using the expression "philosophical bioethics" to maintain the idea that *wisdom* must guide our actions in all these domains and to synthetize the results of particular subdomains into a comprehensive understanding of our moral lives. For instance, it could guide us in discussing neurobioethical issues such as whether we should allow smart drugs to be used in high schools and universities not only for treatment, but also as enhancers.

Before proposing a map to guide us through the landscape of neurobioethics, I would like to clarify better the uses of the neologism "neuroethics." It has, following Roskies' 2002 famous distinction (see references), become commonplace to distinguish two main meanings of "neuroethics", either to refer to: (i) the ethics needed to conduct neuroscientific research and its technological applications (briefly, neuro*ethics*) or (ii) the neuroscientific studies of ethics or, better, morality (briefly, *neuro*ethics). To illustrate, I would like to mention the way the *Global Consortium O3*, in its first meeting, organized the main panels:

1º – Ethics of neuroscience (Chaired by Professor Julian Savulescu, Oxford University);

2º – Neuroscience of ethics (Chaired by Professor Walter Sinnott-Armstrong, Duke University).

In chapter 1 of this book, I will present a case relating to cognitive enhancers and some of the questions on this topic we discussed in the first panel using neurobio*ethics* as the ethics needed for neuroscientific research and some of its neurotechnological applications.

Since I will, in this work, focus on the first group of issues, that is, on the *ethics of neuroscience*, I would here like to mention the main questions the first panel aims to investigate over the next couple of years: (i) the challenges of neurotechnology (e.g., the use of machines to read the human mind, brain–computer interfaces etc.); (ii) the possibility of predicting behavior and controlling individuals (e.g., the use of neurostimulation through tDCS, TMS and DBS); iii) issues surrounding the use of automated systems, especially apps, to increase well-being (e.g., the use of virtual reality to treat mental disorders), etc. I will examine these issues in detail in chapter 3 (cf. www.o3brain.org for further information).

The ethical significance of neuroscientific investigations using human beings as subjects of research or their potential applications is evident. For instance, in the first group, there are conflicts between the increasing accuracy of mind-reading machines against the right to *mental* privacy; in the second, the potential therapeutic benefits of new procedures to treat mental disorders, but, at the same time, the unknown long term effects of, for instance, the use of mood (e.g., lithium) or cognitive enhancers (e.g., amphetamine, commercially known as Adderall®) and issues related to individual identity *and personhood*; in the third, the conflicts between increasing efficacy *versus* potential abuses of technological devices.

Considering these issues, the first panel of the O3 group sorted out some key research problems:

1) How would the development of more powerful neuroimaging that was better equipped to detect consciousness affect the ethics of end-of-life medical decisions?
2) How should decisions to deploy medical interventions which directly affect the human mind be made?
3) Such interventions raise questions about what role, if any, they should have in the criminal justice system, and whether their use would unacceptably threaten human freedom?

4) How is 'care' constituted when it is delivered via an algorithm, or a robot? Are there particular concerns about care delivery in the context of mental health and well-being?
5) How should mental health and well-being apps be regulated?
6) Is there a universal morality that should guide normative judgments about interventions to improve moral behavior and decision-making?

I will give a brief response, with no intention of presenting a final answer, to some of these questions in chapters 1 and 2. In general though, in this book, I will focus more on issues around the ethics of neuroscience. To be even more specific, I will investigate the permissibility of one special kind of *enhancement*, namely the improvement of cognitive capabilities either using conventional (education) or nor-conventional (genetic or pharmacological) means. This issue relates to questions 1, 4 and 5. To deal adequately with neuroenhancement, I will defend *a* CS-Morality comprising care and respect, which are based on practical cognitivism, to respond to question 6 and, to some extent, questions 2 and 3.

Now, the other meaning of "*neuro*ethics", namely the *neuroscience of ethics*, which would perhaps be better referred to as "the neuroscientific studies of moral phenomena" is a very recent topic of research. According to the opening speech by Professor Walter Sinnott-Armstrong in the above mentioned O3 first meeting, "since the first fMRI study of ethics in Brazil in 2001, neuroscientists have published an increasing number of sophisticated experiments on various aspects of moral judgments, decisions, and emotions." (http://o3brain.org/about-o3/). Actually, as far as I know, the first Brazilian fMRI study on the neural correlates of moral judgement was published in *Neurology* (Oliveira-Souza & Moll, 2000) and predates the first neuroethics studies by Joshua Greene (2001). In any case, the works of the Brazilian neuroscientist Jorge Moll, from the IDOR (Instituto D'Or Pesquisa e Ensino, RJ) are certainly pioneering in the field of the neuroscience of moral judgement, as Professor Sinnott-Armstrong correctly pointed out. Now, the neuroscience of morality deals with questions such as whether there is free will, the role of reason and emotions in moral decision-making etc. There is even some hope that neuroscience can help to settle *normative* issues, for instance, whether a deontological approach or a consequentialist one must be assumed in dealing with moral issues regarding what to do. To give an example: recently, Joshua Greene (2013, 289f; 2017,11) was still trying to make the point that neuroscience leads to utilitarianism as "the *ultimate standard*" against which to judge our moral thinking, contrary to deontology or virtue ethics.

I am somewhat skeptical about whether *neuro*ethics can deliver on most of its ambitious promises. In fact, I think that it is even an inappropriate

expression since the neuroscience of morality still is just *science*, not *ethics*. It is a descriptive enterprise, not a normative one. That is to say, strictly speaking a neuro-*scientific* study of morality is scientific only: it presupposes *explanations* of what is going on in our brain when we think about moral issues, but it cannot *prescribe* a particular moral way of thinking about them and acting accordingly. This is one of the main reasons why we need to move in the direction of neuro*bioethics*.

I would then argue for this thesis, namely that the use of "*neuro*ethics" as a neuroscientific study of ethics commits a category mistake by crisscrossing, so to speak, two different language-games (descriptive and normative), and, consequently, it misuses the concepts "neuro" and "ethics." If this is right, then there is only one employment of "neuroethics" that makes sense, namely, an examination of the ethics of neuroscientific studies and their neurotechnological applications (neuro*ethics*). Therefore, we should stick to what was perhaps one of the first definitions of "neuroethics" given by William Safire at the San Francisco inaugural meeting *Neuroethics: Mapping the Field*, back in 2002: "*the examination of what is right and wrong, good and bad about the treatment of, perfection of, or unwelcome invasion of and worrisome manipulation of the human brain*." (https://danablog.org). I believe this is one of the best definitions of "neuroethics" put forward up to now, and it reveals clear interconnections with traditional bioethics.

Suppose, however, one disagrees with my claim and would like to insist on using the word "neuroethics" in such a way that includes the neuroscientific studies of morality. Well, I will not fight over words or labels here, but I would rather ask why we should attribute any special status to neuroscience anyway. After all, a neuroimage of the brain activity is insufficient to determine whether we should allow, for instance, active, nonvoluntary euthanasia. Besides, I would object that a *scientific* study of morality is certainly relevant, but then I would invite my adversary to think of whether ethnology (or anthropology), sociobiology, primatology, etc. are not as crucial as neuroscience. We cannot certainly hope to use fMRI to look for particular rights or obligations in our brains since they are historical and social constructions, even if they are justified using natural laws. Furthermore, neuroethicists are making an interesting move these days to overcome a unilateral approach, namely they are looking at society as a whole to better understand what is going on in our brains when we make moral decisions. For these reasons, I would stand by the single use of "neuro*ethics*" as the ethics of neuroscience. This is, at least, the main use in this work of "neuro*ethics*", and I will present further arguments later for rehaping it as neuro*bio*ethics.

This restriction brings us to the task of deepening the analysis of the interfaces between neuroethics so conceived, and bioethics. I would then like to quote here the predominant view on this point given by one of the participants of the O3, namely the psychologist Marta J. Farah: "New ethical issues are arising as neuroscience gives us unprecedented ways to understand *the human mind and to predict, influence, and even control it*. These issues lead us beyond the boundaries of bioethics into the philosophy of mind, psychology, theology, law and neuroscience itself. It is this larger set of issues that has…earned it a name of its own." (Farah 2010, 2; italics added) There are two main ideas here: one is the need for interdisciplinarity; the other, the idea that neuroscientific knowledge leads to new technological applications and, consequently, gives us ways to manipulate the brain to control the mind. The first point is common to many studies, including bioethics. However, the second issue just reveals that the proper domain of neuroethics is *the ethics of neuroscience* and its neurotechnological applications. Therefore, to the extent that neuroethicists deal with questions on mental disorders (panic, anxiety, depression, psychosis, schizophrenia etc.), physical brain impairments or addiction, autism, neuroenhancement or treatment, etc. in *scientific research* or *clinical* practices, they seem to fall into the traditional domain of bioethics.

Once again, perhaps the critic could easily respond to the question *How is neuroethics different from bioethics?* in terms of neuroexceptionalism (Illes 2017, ix). That is to say, unlike other organs, the brain connects each of us to who we are; that is, the brain is very special because it links us to our identity and it is the organ of human mind. There are many interesting facts about our brain, but does it pose a new normative or evaluative challenge? No doubt, it has a very special role, for instance, in our capacity to make decisions, in our autonomy, and so on. But the *ethical* question is: should we respect autonomous decisions? Then, it does not follow that it makes sense, as we will see, to speak in terms of "the moral brain." The person, not one of her organs, is the moral entity to be respected. Rightly, Bennett and Hacker (2003, p.68) called this mistake a "mereological fallacy." In other words, mereological fallacies (taking *parts* such as our brains as if they were the *whole* of an organic being) have no place in a sound neurobioethics; that is, we must know-how to *respect persons* as whole biopsychosocial entities, not only their nervous systems or their brains.

Is there really something distinctive to the ethics of neuroscience which justifies the use of "a new name for a new field?" According to Steven E. Hyman, "The major justification for demarcating neuroethics from the broader field of bioethics derives from the special status of the

brain (Roskies, 2002), which is the causal underpinning of our conscious mental lives and of our behavior." (2013, 96). Once again, however, it does look reductionist to consider only the brain: even if it is special for many reasons (billions of neurons making trillhons of neural connections), it seems that we should regard the human being as a whole, as a unity. We must avoid what Bakhurst (2008, 415) calls "brainism," or the view (i) that an individual's mental life is constituted by states, events and processes in her brain, and (ii) that psychological attributes may legitimately be ascribed to the brain only. We must rather consider the person, not the brain, as a whole in relation to nature and society. Granted, controlling the brain is one way of controlling the mind, but this is not the whole story. If it were, neuroethics would be justified because neuroscience really has given us unprecedented ways "to predict, influence, and even to control [the mind]". (Farah 2010, 2). This may well be so, but the real ethical question is under what conditions this power is justified? My quick answer is this: if we are dealing with mental disorders in an experimental setting, then it seems clear that, for instance in the US, the Belmont principles (respect for persons, beneficence and justice) fully apply; if we are discussing other issues, for instance, around the political control of another's mind, then there are legitime or illegitimate ways of doing it, but caring and respecting persons seems always to be morally required. That is to say, respecting persons entails not manipulating their minds, especially if they are part of vulnerable populations, for instance, prisoners or individuals with mental disorders. I will return to the issues of mental disorders in the last chapter, but it seems clear that caring respect requires working to increasea person's autonomy, not enslaving her.

Bearing all these points in mind, I can, finally, sort out the main interconnections between bioethics and neuroethics and arrive at the proper domain of what I have called "neurobio*ethics*." Actually, there is already a vague use of this expression since it seems clear that some scientific developments in neuroscience raise bioethical questions. Concerns in the area of neurobioethics are put in these terms:

> Some of these debates include whether particular interventions represent treatment or enhancement, the use or misuse of neurogenetic and neuroimaging data, moral ambiguities surrounding the use of neural allo- and xenotransplantation, and ethical dilemmas arising within the contexts of neuroscientific policy, clinical practice and the scope and tenor of current and future research directions.
> (http://www.unescobiochair.org/neurobioethics/).

As can be seen, neurobioethics does not rely exclusively on the neuroscience of ethics. It also incorporates the social dimensions of doing research on human beings as, for instance, the details of the infamous Tukesggee case remind us. Granted, neuroscience poses new ethical challenges, but the main issue at stake is how we can deal with them and neurobioethics has the right normative framework, for instance, the principle respect for persons.

It is also worth remembering that the *National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research*, the publisher of the Belmont Report which sorted out the principle of respect for *persons,* was very interdisciplinary. Another difference, then, between *neuro*ehtics (driven by neuroscientific investigations of the moral brain) and neurobio*ethics* is that the latter is more interdisciplinary. It encompasses contributions from moral psychology, anthropology, ethnology, sociology, etc. *and* society, while tradional neuroethics is limited to *one* science alone: the neuroscience of morality.

Another huge difference is that neurobioethics does not make inadaquete conceptual and factual transpositions from the neuroscientific investigations directly into ethics. For instance, it may well be a neuroscientific "conclusion" that personhood is an illusion rooted in the brain (Farah and Heberlein 2010, p.321), but it does not follow that persons do not exist or that this concept cannot be constructed from a more interdisciplinary perspective, namely from a social and juridical stance. I will return to this point later when presenting an argument against Farah's conclusion. For these and other reasons, it is better to work with neurobioethics and not with traditional neuroethics. Consequently, one of the most fundamental questions I would like to ask in this book is whether or not special bioethics is normatively well equipped to treat all the new ethical challenges posed by neuroscience.

Whether neuroethics is just part of bioethics is, then, not a new debate. Some thinkers hold that it is not, while others maintain that these problems can indeed be dealt with within the traditional bioethical framework, for instance by using the basic norms of the *Belmont Report* (especially respect for persons, a concept some neuroethicists think is an illusion), which will be commented on later, or some variant of these such as principlism, autonomism, etc. The Belmont normative framework may be sufficient to cope with the challenges presented by neuroscience and neurotechnology, and if this is the case, as I believe it is, and there are no other neuroethical issues left out, then neuro*ethics* (in the sense of an"ethics of neuroscience") turns out to be just a subfield of bioethics: *neurobioethics*.

The main normative problem this work addresses is consequently as follows: which are the basic norms needed to deal with the ethical

challenges arising from neuroscience and neurotechnology? Thus, another thesis I will argue for is this one: while we are doing neuroscientific research, neuroexperiments and applying neurotechnologies using human subjects, the bioethical normative framework works well, but we need also to get the right one. That is to say, in chapter 2 I will support the *The Belmont Guidelines*, and criticize other normative approaches derived from them such as principlism, autonomism, etc. I will also in this chapter criticize the main attempts of some neuroethicists, such as Greene and Farah, to establish utilitarianism as the normative ethics for dealing with the problems of the ethics of neuroscience. Thus, I will, after correcting the metaethical problems of bioethics in chapter 1, find a way of integrating care and respect into an unified normative approach. Consequently, I will use the expression "neurobioethics" to refer to a particular sub-field of bioethics assuming the *onus probandi* that this hypothesis can be demonstrated; that is, that traditional bioethicists got the moral principles of this domain right. That is to say, well-understood (neuro)bioethics has the necessary normative resources to deal with the main ethical challenges neuroscience and neurotechnology have presented so far. What is further needed is just willingness to follow principles such as respect for persons in our caring practices and scientific research. Briefly, what is needed is *knowing-how to respect neuroscientific subjects*, namely in a caring way. This is one of the main ideas of this work and is presented under the label "neurobioethics."

This last point reveals also the main justification for this book. I would like to put it in this way: in her excellent work "Neuroscience and Neuroethics in the 21th Century", the psychologist Martha J. Fahar (2013, 761) reviews the main neuroscientific challenges facing us nowadays, especially the enhancement of mental functioning, neuroscience-based education, neuromarketing, the new roles of criminal justice and the law, etc. and presents three new ones: (i) neuroliteracy for the neurocentury, (ii) ownership and control of neurotechnology and (iii) avoiding nihilism. The spirit of this book can then be seen as a response to the third "new" challenge: how can we prevent neuroscience's physicalist explanations of human behavior lapsing into nihilism? My view, then, is that persons are *not* mere physical objects and must be respected in a caring manner as moral entities while we are doing applied ethics, especially neurobioethics. This will be illustrated in chapter 3 when I discuss the permissibility of cognitive enhancement. In fact, I will present an argument to refute the idea that personhood is just an illusion. This will make it possible to work with the principle of respect for persons, to defend a person-centered healthcare system, to put the person before the patient or the subject of neuroscientific research or the use of neurotechnologies, etc. To accomplish this goal, my

main proposal is to redirect neuroethics, freeing it from its naturalistic and utilitarian assumptions, towards neurobio*ethics*.