Sigrid Müller, Piotr Jan Morciniec, Nenad Polgar (Hg.)

Neurowissenschaften in der Diskussion

Anregungen zum Weiterdenken

The Moral Brain

Stimuli for Further Reflection

Bioethik in der Diskussion

Bioethics in discussion

Sigrid Müller, Piotr Jan Morciniec, Nenad Polgar (Hg.)

Neurowissenschaften in der Diskussion

The Moral Brain

facultas

Bioethik in der Diskussion

Im Auftrag der Association of Bioethicists in Central Europe
herausgegeben von Sigrid Müller, Piotr Jan Morciniec und Nenad Polgar

German Proofreading: Claudia Bernal Díaz, Bakk.
English Proofreading: Nenad Polgar, MA Ph.D.
DTP: Jerzy Bosowski

Bibliografische Information der Deutschen Nationalbibliothek
Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.d-nb.de abrufbar.

Permalink (Open Access: Zugang ab August 2023):
http://phaidra.univie.ac.at/o:1119070
Druck: Facultas Verlags- und Buchhandels AG
Printed in Austria
ISBN 978-3-7089-2176-1
DOI: 10.24989/BCE.neuro

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 5–6
DOI: 10.24989/BCE.neuro.9

Content / Inhalt

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 7–10
DOI: 10.24989/BCE.neuro.0

The Moral Brain. Editors' Preface

The association "Bioethicists in Central Europe" (BCE) held its annual meeting in Olomouc on October 22-23, 2021. The focus was on the question of how much morality is possible for human beings. Morality – here understood in the sense of the ability to align one's own actions with personally responsible moral principles – seemed to be sensitively called into question by reports from the neurosciences. Were research results to be trusted which stated that the biological processes of the human being do not only concern the functioning of the organs, but also his moral behaviour? Was freedom, which in the theological and philosophical tradition plays a fundamental role in the moral demands on man, to such a degree conditioned by biological processes that it could only be assumed in a limited way or possibly not at all? Do we humans even resemble the animals we describe in their behaviour, as if they were entirely controlled by stimuli and reactions to them?

The aim of the above-mentioned conference was to investigate these challenges and to discuss the topic from philosophical, theological-ethical, neuroscientific and psychiatric perspectives. The different approaches made it clear that the human brain with its neuronal network is so complex and flexible that one can by no means speak of a one-sided determination of moral action by the brain. Human thinking, interaction, but also practice and new influences can thicken, dilute and supplement the neuronal strings of the piano of morality. However, because of the cross-linking of neurons, the picture of the strings of the piano is woefully inadequate. Another difficulty is that neurological studies are usually very specialized and therefore do not allow us to make general statements about the whole of human behaviour. Ethicists, on the other hand, can formulate inquiries

about generalizations of scientific results and thus contribute to inter- and transdisciplinary dialogue.

We hope that the contributions to this volume will make the complexity of the questions clear, but also encourage critical inquiry and further thinking.

Vienna, June 2022
Sigrid Müller, Nenad Polgar and Piotr Jan Morciniec

Das moralische Gehirn. Ein Vorwort

Die Vereinigung „Bioethicists in Central Europe“ (BCE) hielt von 22. bis 23. Oktober 2021 in Olmütz ihre Jahrestagung ab. Im Mittelpunkt stand die Frage, wieviel Moralität dem Menschen möglich ist. Moralität – hier verstanden im Sinne der Fähigkeit, das eigene Handeln an persönlich verantworteten moralischen Grundsätzen auszurichten – schien durch Berichte aus den Neurowissenschaften empfindlich in Frage gestellt. War Forschungsergebnissen zu trauen, die besagen, dass die biologischen Vollzüge des Menschen nicht nur das Funktionieren der Organe, sondern auch sein moralisches Verhalten betreffen? War die Freiheit, die in der theologischen und philosophischen Tradition eine grundlegende Rolle für die moralischen Ansprüche an den Menschen darstellt, zu einem Grad durch biologische Prozesse bedingt, dass sie nur eingeschränkt oder eventuell gar nicht mehr vorausgesetzt werden durfte? Ähneln wir Menschen gar den Tieren, die wir in ihrem Verhalten beschreiben, als seien sie ganz durch Reize und Reaktionen darauf gesteuert?

Diesen Herausforderungen nachzugehen war das Ziel der genannten Tagung bei dem das Thema aus philosophischer, theologisch-ethischer, neurowissenschaftlicher und psychiatrischer Perspektive diskutiert wurde. Die unterschiedlichen Zugänge ließen erkennen, dass das Gehirn des Menschen mit seinem neuronalen Netzwerk so komplex und flexibel ist, dass man keinesfalls von einer einseitigen Bestimmung des moralischen Handelns durch das Gehirn sprechen kann. Menschliches Denken, Interaktion, aber auch Übung und neue Einflüsse können die neuronalen Saiten des Klaviers der Moralität verdicken, verdünnen und ergänzen. Aufgrund der Quervernetzungen der Neuronen ist das Bild von den Saiten des Klaviers aber leider unzureichend. Eine weitere Schwierigkeit besteht dar-

in, dass neurologische Untersuchungen in der Regel sehr spezialisiert sind und daher nicht erlauben, generelle Aussagen über das gesamte Verhalten des Menschen zu machen. Ethiker*innen wiederum können Anfragen an Verallgemeinerungen naturwissenschaftlicher Ergebnisse formulieren und so zum inter- und transdisziplinären Dialog beitragen.

Wir hoffen, dass die Beiträge dieses Bandes die Komplexität der Fragen deutlich werden lassen, aber auch Mut machen zum kritischen Nachfragen und Weiterdenken.

Wien, im Juni 2022
Sigrid Müller, Nenad Polgar und Piotr Jan Morciniec

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 11–24
DOI: 10.24989/BCE.neuro.1

The Brain and Morality. Introductory Remarks

Das moralische Gehirn. Zur Einführung

Nenad Polgar, Vienna,
Sigrid Müller, Vienna,
Piotr Jan Morciniec, Opole

Abstract (Deutsch):

Dieser Artikel führt in die Thematik des Buchbandes ein, indem er die verschiedenen Denkansätze zum Thema Moralität und Gehirn aufzeigt. Dabei gehen die Autor*innen davon aus, dass der Mensch ein moralisches Wesen ist und dem Thema Moral in diesem Zusammenhang ein eigener Bereich zusteht. In der Untersuchung von Moralität und Gehirn spielen die Neurowissenschaften eine wichtige Rolle. Aber nicht nur diese sind von Relevanz, sondern es ist ein interdisziplinärer Dialog mit verschiedenen Geisteswissenschaften, darunter auch der Ethik, notwendig, woraus sich der Bereich der Neuroethik entwickeln konnte. Das Hauptaugenmerk liegt dabei auf den ethischen Auswirkungen der Neurowissenschaften und deren Bewertung. Bei der Zurechnung von moralischen Handlungen wird die immense Komplexität, Plastizität und Vernetzung der verschiedenen Gehirnbereiche beachtet und Aussagen zur Determinierung getroffen. In diesem Zusammenhang stellt sich die Frage nach der Freiheit des Menschen. Ein letzter Bereich, welchen dieser Artikel anspricht, ist die Verbesserung des Gehirns, wobei festgestellt wird, dass ein kognitives Enhancement nicht unbedingt zu einer moralischen Anwendung der gesteigerten Erkenntnisfähigkeit beiträgt und damit eher gegen ein solches zu argumentieren ist.

Abstract (English):

This article introduces the topic of the book volume by showing the different approaches to the topic of morality and the brain. The authors assume that human beings are moral beings and that the topic of morality has its own area in this context. Neuroscience plays an important role in the study of morality and the brain. However, not only these are relevant, but an interdisciplinary dialogue with various humanities, including ethics, is necessary, from which the field of neuroethics could develop. The main focus is on the ethical implications of neuroscience and their evaluation. In the attribution of moral actions, the immense complexity, plasticity and interconnectedness of the different brain areas are taken into account and statements on determinism are made. In this context, the question of human freedom arises. A last area, which this article addresses, is the enhancement of the brain, whereby it is stated that a cognitive enhancement does not necessarily contribute to a moral application of the increased cognitive ability and thus it is rather to be argued against such an enhancement.

Keywords (Deutsch):
Gehirn; Neurowissenschaften; Neuroethik; Determinierung; Enhancement; Freiheit; Moralität; Ethik

Keywords (English):
brain; neuroscience; neuroethics; determinism; enhancement; freedom; morality; ethics

Human being are moral beings. They possess the faculties of reason and will that allow them not only to make judgements about right and wrong, but also to follow through on those judgements by making specifically moral decisions and, consequently, being morally accountable. Although this kind of decisions might and regularly do overlap with other kinds of decisions – e.g., medical, legal, social, practical – there are strong reasons to believe that morality is not reducible to these or any other area of human life, but instead constitutes a distinct area on its own. That conviction runs so deep that

one might be entirely justified in arguing that human beings possess moral brains. Or so we thought until very recently.

1. Do we have moral brains? The input of neurosciences on ethical reflection

Since neurosciences started to examine the different parts of the brain and to measure its activity during different actions, doubts were raised as to whether human beings possess an area which could be called moral brains. Also how the measured activity needs to be interpreted, has been fiercely discussed between neuroscientists and experts from other academic fields in the last few decades.

In one sense, that discussion is not new. Serious philosophical accounts (for instance, David Hume's or Alfred Ayer's) have disputed specificity of moral judgements for centuries and, more recently, this kind of accounts found an unexpected ally in sociobiological research aimed at portraying morality as a mere evolutionary adaptation or mechanism (Clayton & Schloss 2004). Nevertheless, for a variety of reasons, these and similar attempts at discrediting morality as a distinct area of human life never acquired a large following and were mostly treated as a source of peculiar and largely technical academic debates that were of interest to a very limited circle of experts. As opposed to that, neurosciences seem to pose a new challenge within this long-running discussion on the nature of morality, but before this could be appreciated a few words need to be said about what neurosciences are.

Neuroscience is a fairly extensive and rapidly developing area of research, combining multiple disciplines (physiology, anatomy, molecular biology, cytology, computer sciences, etc.) in order to study the nervous system and the brain in terms of their structure, working, development, and malfunctioning (Amthor *et al.* 2020; Felten *et al.* 2022). Due to the fast development of relevant technologies, neuroscience acquired new tools (fMRI, PET, SPECT, EEG, MEG, etc.)[1] for probing deeper into

[1] The abbreviations refer to Functional magnetic resonance imaging, positron emission tomography, Single photon emission computed tomography, Electroencephalography, Magnetoencephalography).

the brain and managed to perfect its experimental techniques and methods of investigation of this most complex of human organs in the recent decades.[2] The potential of that research is staggering in its implications, insofar as it not only promises an unprecedented understanding of neural mechanisms, but also a major step forward in treating and managing brain-related diseases, development of new brain-altering technologies, and possibly influencing some of our long-held beliefs about the very nature of human beings.

2. The raise of neuroethics

Through all of these implications, neuroscience crosses paths with humanities and other fields of disciplines in general and ethics in particular, raising what seems to be a host of new issues[3] and urging an interdisciplinary dialogue. In the last few decades, that dialogue has been ongoing within the bounds of a new discipline called neuroethics.[4] The importance of the development of this new discipline has been compared by some researchers with the development of bioethics during the previous generation (Levy 2007, ix–x). Just like bioethics, neuroethics is in its core interdisciplinary, they both developed explosively at the time of their initiation, and they both grapple with issues that are, in the very literal sense, of vital interest to us as human beings. As Neil Messer argues:

> Whatever view one takes of the relationship between brain, mind and soul, it is evident that our brains play a distinctive and essential part in more

[2] For an introduction into the functioning of the transmission of information in the brain see Gazzaniga *et al.* (2014), 22–69 (ch. 2).

[3] The scope of what is being discussed at the crossroads of neuroethics is too extensive to even list in this article. A good overview is, however, provided in the following four-part bibliography, composed by a group of researchers (Buniak *et al.* 2014; Darragh *et al.* 2015; Martin *et al.* 2016; Becker *et al.* 2017).

[4] Alongside neuroethics, the interdisciplinary dialogue between neurosciences and other disciplines gave birth to a host of other new disciplines or research areas like neuroanthropology, neuroeconomics, neuroeducation, neurosociology, etc. (Leefmann & Hildt 2018, 15).

> or less every aspect of our lives as human persons, from the regulation of the basic functions that keep us alive to emotions and desires, thinking, willing and acting, understanding and beliefs about the world, self-understanding and personal identity, personal and social relationships, and even our experience and practice of faith. (Messer, 2017, 1–2)

One of the pioneers of the new discipline, Adina Roskies, divides neuroethics into two distinct parts; the ethics of neuroscience and the neuroscience of ethics (Roskies 2002, 21–23). The first, less significant, part of neuroethics – also called "ethics of practice" – deals with ethical issues related to the execution of the kind of research done in neuroscience and is of little interest to us in this article, since its concerns are largely covered already by the older discipline of bioethics. The second part of neuroethics, however, engages with ethical implications of neuroscience and the evaluation of impact results of studies in this discipline might have on existing social, ethical, and legal structures. This part of neuroethics, therefore, covers the kind of issues that emerge as neuroscience probes deeper into the biological mechanisms of the brain that underlie and, therefore, (co-)determine our lives as human persons.

3. The physiology of the brain and the imputation of moral acts

To reflect this relationship between brain and morality may lead to a one-directional view of the brain presenting the physical condition for moral actions. In such a line of thought, one could argue that it is the individual brain which determines the moral capability of a person and thus shapes, enables or minimizes individual freedom and responsibility. Ultimately, this would lead to physical examination of the brain in the context of moral and criminal judgment over moral misbehaviour of human persons, or – if an entirely deterministic model is applied – to interpreting all misbehaviour as caused by the brain alone, which would of course result

in an entire loss of significance of moral and legal notions as culpability, guilt and repentance.[5]

To support such a one-directional view, one could point at "traditional" examples in the history of neurosciences that were used to point at the causative effect of the brain on moral behaviour, as the famous example of Phineas Gage whose brain was injured with the effect that the brain lesion lead to a drastic change of his social capabilities (Spezio 2011, 344–345). More cautious interpretations would probably say that it is difficult to argue that the effects of such a massive injury allow concluding more than the fact that a relationship exists between the physical shape of the brain and human action. What the example can show is how the lack of certain brain areas can provide an obstacle for certain actions, but it does not necessarily allow to conclude that moral acts are exclusively shaped by the physiology of the brain. For example, if we look at neighbour disciplines as psychotherapy and psychiatry, we can learn from their interventions that human beings are far from being entirely determined. Rather, reflection and asking the right questions can provoke changes in behaviour, and multiply behavioural options.[6] This is due to the immense complexity, plasticity and interconnectedness of different brain areas.

The question of the location of the moral brain has also been raised and the view that all moral actions are determined by one specific part of the brain has proofed to be erroneous. For example, different operations in the brain are often interconnected, as more detailed studies show. Rational thinking, emotional reactions, and social awareness are usually linked to one another and cannot be separated: "what is emerging is a complex interconnection of circuits in which emotional signals cannot be separated from adaptive reasoning and decision making when such judgment and action are relevant for oneself and others" (Spezio 2011, 352).

Other findings in the context of psychological theory correspond to this interconnectedness. For example, it has become clear that the best moral decisions are made if different levels of experience and emotions

[5] Cf. the contribution by Petr Hluštík in this volume.

[6] Cf. Borut Škodlar, *Psychotic Discorders and Personal Freedom*, in this volume.

comply with moral reasoning (Grill 2021). However, a decision in which rational and emotional reactions are in entire consonance presupposes a high level of self-control in the human being (Bauer 2018, 211). The interconnectedness of different brain areas also explains why not only the brain shapes moral behaviour, but rather practice, will and exterior stimuli can also shape the "moral brain". Therefore, recently neuroscientists have highlighted the importance of good forms of education as important ways of modelling the brains in a way that helps people to pursue their happiness (Bauer 2018, 56).

4. Brain activity and moral decision making

A broad discussion on whether the brain makes moral decisions – rather than the human self or "I" – arose after Benjamin Libet's famous experiments in neuroscience in which he asked subjects to move their finger whenever they wanted to, while the experiment recorded their readiness potential to do so in the brain. The experiment showed that such readiness appeared and was recorded in the brain about 400 milliseconds before the conscious intention to move the finger was formed (Libet *et al.* 1983). Thus, some conclude that humans are not consciously making such or any other decisions, which is why one cannot blame them (i.e. their conscious selves) for morally problematic behaviour.

While the experiment is certainly intriguing and demonstrates the potential of neuroscience not only to expand our knowledge of how the brain works, but also to challenge other disciplines to take those insights into account, one has to remain cautious not to read into it more than such and similar experiments could possibly settle. In other words, the interpretation that the experiment showed that there is no free will because the conscious self is not the primary agent who makes decisions, is not the only credible reading of the results of the study. Other researchers and philosophers proposed alternative interpretations, ranging from a proposal that our consciousness initiates regularly only important decisions and leaves the details to subpersonal processes, to arguing that the experiment still leaves plenty of space for certain views of moral responsibility (cf. Deecke

2008), to questioning some of the assumptions about intention and consciousness Libet made in the experiment and, consequently, what he actually measured (Levy 2007, 227–229). While a closer analysis of the details of the experiment and alternative interpretations of its results[7] would take us too far off the tracks in this contribution, the main point we wanted to make is that whether there is free will is too complex an issue to be settled by relying exclusively on neuroscientific experiments.

5. Positions in the discussion about mind and brain

In the discussion on the relationship between the physical brain and the human activity of the mind, different positions are taken, which usually oscillate between complete determinism (all activities of the mind can be explained by activities in the brain) and non-determinism (mental acts exist which are "not part of the physical or biological structures"[8]). Others intend to save aspects of the discourse on freedom and to accept determinism. In this way, Seidel distinguishes between "free will" in a metaphysical sense and "personal will" ("eigener Wille"). While he declines the existence of a metaphysical, absolute free will and accepts determinism which he understands as the fact that the brain is a causative factor of actions, he defends, at the same time, practical freedom and responsibility. Both are possible and necessary due to the immense amount of different options provided within the framework of the functioning of the brain (Seidel 2009, 186–187). Freedom, in this sense, is a functional setting of an evolutionary perspective of the development of the human species, by which moral actions should be sanctioned and shaped for the greater good of the human species. Even if brain activity has proofed to be fun-

[7] Bauer 2018, Fn. 17, 194–202, provides an overview over the criticism of B. Libet himself and other neuroscientists of the interpretation of Libet's experiment by G. Roth and W. Singer at the level of the technical setup of the experiment (exactness), the interpretation of the slow cortical potential in this context (origin) and the question whether the activity measured in the experiment is the free act itself or its putting into action, which would inverse the results.

[8] W. Schaupp (see contribution to this volume).

damental, a huge variety of choices exists which can be interpreted in different ways as providing certain degrees of freedom. Freedom thus is limited to the freedom of choice; a transcendental freedom as constitutive feature of the human being cannot be considered in such a framework (Seidel 2009, 23). In a way, there is no further responsibility but for the survival of the species.

While Seidel intends to bridge the perspective of natural sciences with the approach to the question of human freedom in the field of Humanities by explaining the results of neurosciences in such a way that they provide space for what human beings experience as freedom, Rosenberger argues for the legitimate existence of both paradigms (determinism and freedom) next to each other because of epistemological reasons. This approach has been called the "double aspect theory"[9]. Since every scientific approach is confined to a method, the constructivism attached to each method provokes that different epistemological systems operate that cannot be reduced to each other. According to Rosenberger, we need to get acquainted to the fact that freedom, including transcendental freedom, is linked to the first-person perspective examined in humanities while determinism is a concept related to the third-person perspective which is the perspective of natural sciences (Rosenberger 2006, 228–230). While the first-person perspective looks at reasons, the third-person perspective looks at causes. The strength of the "double aspect theory" is that it allows to justify from a methodological point of view that the whole range of philosophical notions of freedom including fundamental freedom can be maintained.[10]

Returning back to the beginning of this article and the issue of whether human beings possess moral brains, one might, therefore, argue that a significant part of the discussion in the neuroscience of ethics hinges on whether one should include the prefix "co-" into the description of the determining role of brain mechanisms on our lives. Therefore, the morally

[9] See W. Schaupp's contribution in this volume.

[10] A question related to this position (and that of some forms of determinism) is that the methodological reductionism of determinism can be applied only to actions in the past, because the decision making process is open to influencing factors.

decisive question regarding neurosciences deals with the inter-relationship of physicality and moral action, which is basically the right balance between deterministic visions of the human being and those emphasizing freedom of decision making, freedom of the will, even if this freedom is not absolute, but partially based on its "natural", neuro-physical basis. Which position we take among the variety of deterministic, weak deterministic and non-deterministic visions of this relationship has experimentally been related also to personal character traits (Feltz and Cokely, 306). The challenge of bridging neurosciences and philosophical concepts of the "self" has been discussed at least since 2004 when 11 neuroscientists published their *Manifesto about Present and Future of Brain Research*[11] in 2004 and showed that thorough and ongoing philosophical reflection is needed to avoid reductionist models of interpretation (Langthaler 2008, 68). Therefore, discussions are far from coming to an end.

6. Brain activity and moral theory

This, of course, does not mean that such experiments and neuroscientific research in general have no bearing on our understanding of the moral brain and moral responsibility. As J. Leefmann and E. Hildt have argued, neuroscientific research has a tremendous potential to reveal cognitive and normative biases that affect agents in making moral decisions and, thus, not only reveal when the demand for moral conduct is misplaced, but might also have an impact on further development of ethical theories, insofar as these presuppose a certain view of a moral agent (Leefmann & Hildt 2018, 17–18).

A good example of such potential impact of neuroscientific research on ethics are experiments conducted by Joshua Greene in which he measured activity of various brain regions by using fMRI as his subjects grappled with various non-moral and moral dilemmas (Greene *et al.* 2001). The experiments led Greene to conclude that human beings utilise two kinds of moral reasoning and decision-making; an intuitive-automatic de-

[11] Cf. the contribution by Angelika Walser in this volume.

cision-making, associated with higher activity in those regions of the brain that are responsible for emotional processing and social cognition, and a slow-going, rational decision-making, associated with higher activity in those regions of the brain responsible for problem-solving and cognitive processing. On the basis of these results, Greene argued that the intuitive-automatic decision-making is reflected in deontological ethical theories and moral reasoning, while the slow-going, rational decision-making is reflected in utilitarian ethical theories and moral reasoning (Greene 2014, 132–143).

Greene has been criticised heavily on various aspects of his studies (Messer 2017, 39–59) – for instance, on his identification of morality with altruism, his assumptions about deontological and utilitarian ethical theories, etc. Nevertheless, one has to take into account that the kind of research Greene engaged in is still in a very early phase and might be improved upon in the future. Furthermore, since the normative-ethical conclusions Greene arrived at on the basis of his research – albeit still problematic and unwarranted, given the current design of his studies and their assumptions – are more modest than those proceeding from Libet's studies, they are also much more likely to inform and influence the debates in neuroethics and have a lasting impact on ethics in general.

7. Improving morality by enhancing the brain?

In spite of the different explanations of the relationship between mind and brain, the importance of morality in society is usually not denied. Rather, the insights of neurosciences are proposed to be applied for improving moral behaviour, or for expanding spaces of freedom of the individual. Such a development can also lead to the question of neuroenhancement with regard to morality. This is a complex issue and cannot be located in one "moral area" of the brain (cf. Banja 2018, 300). Rather, it involves cognitive and emotional activities which are located in different areas of the brain. Hormons that influence emotions, as serotonin and oxytocin, cannot responsibly be used only where there they are lacking in comparison to "normal" production in the body. To use them for enhance-

ment could produce unintended effects. Similarly, cognitive enhancement would not necessarily contribute to a moral application of the enhanced ability for cognition. Therefore, arguments rather speak against moral enhancement (Fenner 2019, 234–239). Related to this, also the need for specifying human rights to protect human beings from certain inventions in the brain has been spelled out (Ienca & Andorno 2017).

There are many more questions related to the agenda of neuroethics in terms of neuroscience of ethics which we are unable to touch in this short contribution. But even the very brief points that were mentioned show that neuroethics promises to continue provoking ethical thought in the next couple of years and decades.[12]

Bibliography

Amthor, F. R. *et al.* (eds), Essentials of Modern Neuroscience, New York, NY 2020.

Banja, J. B., Moral Reasoning, in: L. S. M. Johnson, K. S. Rommelfanger (eds), The Routledge Handbook of Neuroethics, New York, NY 2018, 287–303.

Bauer, J., Selbststeuerung. Die Wiederentdeckung des freien Willens, München 2018 [2015].

Becker, K. *et al.*, A Four-Part Working Bibliography of Neuroethics, Part 4. Ethical Issues in Clinical and Social Applications of Neuroscience, "Philosophy, Ethics, and Humanities in Medicine", 2017, 12, 1, https://www.researchgate.net/publication/317319078_A_four-part_ working_bibliography_of_neuroethics_Part_4_-_Ethical_issues_in_clinical_and_social_applications_of_neuroscience.

Buniak L. *et al.*, A Four-Part Working Bibliography of Neuroethics, Part 1. Overview and Reviews – Defining and Describing the Field and Its Practices, "Philosophy, Ethics, and Humanities in Medicine", 2014, 9, 1, https://www.researchgate.net/publication/262781235 _A _four-part_working_bibliography_of_neuroethics_Part_1_Overview_and_reviews_defining_and_describing_the_field_and_its_practices.

[12] Cf. the contribution by Andrea Vicini in this volume.

Clayton, P., Schloss, J. (eds), Evolution and Ethics. Human Morality in Biological and Religious Perspective, Grand Rapids, MI 2004.

Darragh, M. *et al.*, A Four-Part Working Bibliography of Neuroethics, Part 2. Neuroscientific Studies of Morality and Ethics, "Philosophy, Ethics, and Humanities in Medicine", 2015, 10, 1, https://www.researchgate.net/publication/273519033_A_four-part_working_bibliography_ of_neuroethics_part_2--Neuroscientific_studies_of_morality_and_ethics.

Deecke, L., Was ist Geist aus der Sicht der Hirnforschung?, in: K. Appel *et al.* (eds), Naturalisierung des Geistes? Beiträge zur gegenwärtigen Debatte um den Geist, Würzburg 2008, 92–134.

Felten, D. L. *et al.* (eds), Netter's Atlas of Neuroscience, Philadelphia, PA 2022.

Feltz, A., Cokely, E.T., Informing ethical decision making, in: L. S. M. Johnson, K.S. Rommelfanger (eds.), The Routledge Handbook of Neuroethics, New York, NY 2018, 304–318.

Fenner, D., Selbstoptimierung und Enhancement. Ein ethischer Grundriss, Tübingen 2019.

Gazzaniga, M. S., Ivry, R. B., Mangun, G. R., Cognitive Neurosicence: The Biology of the Mind, New York, N.Y. [4]2014.

Greene, J. D. *et al.*, An fMRI Investigation of Emotional Engagement in Moral Judgment, "Science", 2001, 293, 5537, 2105–2108.

Greene, J. D., Moral Tribes. Emotion, Reason, and the Gap between Us and Them, London 2014.

Grill, R., Willensschwäche. Eine moralpsychologische Untersuchung, (Studien zur theologischen Ethik, 157), Basel/Würzburg 2021.

Ienca, M., Andorno, R., Towards new human rights in the age of neuroscience and neurotechnology, "Life Sciences, Society and Policy", 2017, 13, 5, 1–27, DOI 10.1186/s40504-017-0050-1.

Langthaler, R., Naturalisierung des Geistes? Anmerkungen zu einer dominanten Debatte in der Gegenwartsphilosophie, in: K. Appel *et al.* (eds), Naturalisierung des Geistes? Beiträge zur gegenwärtigen Debatte um den Geist, Würzburg 2008, 12–71.

Leefmann, J., Hildt, E., Neuroethics and the Neuroscientific Turn, in: L. S. M. Johnson, K. S. Rommelfanger (eds.), The Routledge Handbook of Neuroethics, New York, NY 2018, 14–32.

Martin, A. *et al.*, A Four-Part Working Bibliography of Neuroethics, Part 3. 'Second Tradition Neuroethics' – Ethical Issues in Neuroscience, "Philosophy, Ethics, and Humanities in Medicine", 2016, 11, 1, https://www.researchgate.net/publication/308399296_A_four-part_ working_bibliography_of_neuroethics_part_3_-_second_tradition_neuroethics_-_ethical_ issues_in_neuroscience.

Levy, N., Neuroethics. Challenges for the 21st Century, Cambridge 2007.

Libet, B. *et al.*, Time of Conscious Intention to Act in Relation to Onset of Cerebral Activity (Readiness-Potential). The Unconscious Initiation of a Freely Voluntary Act, "Brain", 1983, 106, 3, 623–642.

Messer, N., Theological Neuroethics. Christian Ethics Meets the Science of the Brain, London 2017.

Roskies, A., Neuroethics for the New Millennium, "Neuron", 2002, 35, 1, 21–23.

Seidel, W., Das ethische Gehirn. Der determinierte Wille und die eigene Verantwortung, Heidelberg 2009.

Spezio, M. L., The neuroscience of emotion and reasoning in social contexts: implications for moral theology, "Modern Theology", 2011, 27, 2, 339–356.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 25–44
DOI: 10.24989/BCE.neuro.2

Neuroscience and Theological Bioethics: An Interdisciplinary Dialogue

Neurowissenschaft und theologische Bioethik: ein interdisziplinärer Dialog

Andrea Vicini, Boston

Abstract (Deutsch)
Die theologische Bioethik bezieht sich in vielfältiger Weise auf die Neurowissenschaften: Sie untersucht deren konkrete Anwendungen (z. B. Tiefenhirnstimulation), beschäftigt sich mit den theoretischen Ansätzen, mit welchen Philosophie und Theologie neurowissenschaftliche Themen bearbeiten, sowie mit den Rahmentheorien, die im Hintergrund der Neurowissenschaften stehen. Von den theologischen Zugängen hebt der Autor die persönliche und soziale Dimension von Heilung hervor sowie Debatte über die Unterscheidung zwischen Therapie und Verbesserung des Menschen (Enhancement). Das ethische Ziel ist es, einen umfassenderen ethischen Rahmen zu artikulieren, der es ermöglicht, Dynamiken zu erkennen und persönliche und soziale Tugenden zu fördern, mit Hilfe derer ungerechte soziale Strukturen transformiert werden können. Beiträge aus dem Gebiet der Neurowissenschaften, Philosophie, protestantischen und römisch-katholischen Theologie zielen darauf ab, die theologische Auseinandersetzung mit den Neurowissenschaften weiter anzuregen. Dafür muss kritisch analysiert werden, welche Visionen sowohl für Neurowissenschaftler*innen als auch für theologische Bioethiker*innen auf den drei Ebenen der Person, Wissenschaft und Gesellschaft leitend sind.

Abstract (English)
Theological bioethics engages neuroscience by examining its concrete applications (e.g., deep brain stimulation), the philosophical and theological approaches that are proposed to discuss them, as well as those that frame the whole field of neuroscience. Among the contributions offered by theological discourse, the author highlights reflections on the personal and social dimensions of healing, and debates regarding the distinction between therapies and enhancements. The ethical goal is to articulate a more comprehensive ethical framework that discerns and that aims at fostering personal and social virtuous dynamics capable of transforming unjust social structures. Contributions from neuroscientists, philosophers, Protestant scholars, and Roman Catholic authors aim at further stimulating theological engagements with neuroscience. What emerges is the need for a threefold approach that critically examines the visions of the person, of science, and of society that animate both neuroscientists and theological bioethicists.

Keywords (Deutsch)
Neurowissenschaft; Tiefenhirnstimulation; theologische Bioethik; Heilung; Enhancement; Tugenden;

Keywords (English)
Neuroscience; deep brain stimulation; theological bioethics; healing; enhancement; virtues;

1. Introduction: Deep Brain Stimulation

A recent article published in the prestigious newspaper *New York Times* shared the story of Sarah (Belluck 2021). Because Sarah was affected by a very serious depression that greatly limited her ability to live on her own and work, that resisted any available treatment, and that even led her to attempt suicide, she accepted to undergo an experimental treatment. After studying the electric patterns in her brain, and identifying which patterns were associated to her depression, the research team placed a neurostimu-

lator in her brain, in the area where the electric patterns associated with her crises of depression were located. Differently from other intracranial brain stimulators, which can be regulated by patients or that stimulate the nearby neurons continuously, her device briefly stimulates the brain only when the electric patterns associated with her crises of depression are detected (Scangos, Khambhati *et al.* 2021). The benefits were unexpected and welcomed: she returned to live on her own, resumed her work, enjoyed her life, even laughed again after over five years.

Brain stimulators are not new. In the U.S.A., the Food and Drug Administration, which is responsible for protecting the public health by ensuring the safety, efficacy, and security of medical devices, approved the use of brain stimulators for specific neurological disorders: essential tremor (1996), neurological movements disorders (dystonia, in 2002), severe epilepsy (2019), and Parkinson's disease (1997, 2016, and 2020)[1]. Besides these approved uses, clinicians are experimenting whether various types of deep brain stimulators could be beneficial in addressing the symptoms of selected psychiatric disorders and other neurological diseases currently untreatable that seriously affect patients and their quality of life.

Other narratives could be added to Sarah's story to highlight how, in clinical settings, neuroscience attempts to address the concerns of patients affected by pathologies that, at the moment, cannot be treated successfully in any other way. The goal of researchers and clinicians is to expand the ways in which the symptoms of these patients could be treated, by reducing or eliminating them, promoting their quality of life and well-being, and striving to achieve healing.

From an ethical point of view, while this commitment to care for patients is needed and is praiseworthy, ethical concerns should be identified, named, and addressed. Particularly, research in neuroscience should allow us to understand better what affects our brains and how neurostimulators

[1] In the U.S., the Food and Drug Administration regulates also human and veterinary drugs, biological and biopharmaceutical products, food safety, tobacco products, dietary supplements, vaccines, blood transfusions, electromagnetic radiation emitting devices, and cosmetics, animal foods and feed. In Europe, its analogous is the European Medicines Agency. However, the medical devices are regulated by national competent authorities.

interact with normal brain functioning as well as with specific pathologies. Further clarifications and reassurances are needed regarding the possible damages to neurons that depend on introducing and placing deep brain stimulators (Messer 2017, 109 and 146–147).

In many instances the patients report subjective benefits that greatly improve their personal and social life. However, the descriptions of family members, friends, and colleagues might not match these accounts and highlight unexpected changes in behavior and in the personality of those undergoing deep brain stimulation. Hence, studies are needed to assess, compare, and contrast the outcomes, which might differ according to who is evaluating the benefits or consequences of using deep brain stimulators.

Finally, it seems necessary a national and international accounting and monitoring of these stimulators by examining for which pathologies they are used, whether these applications are experimental or whether the approval of the responsible authorities has been granted, and what surfaces in pursuing follow-up studies (Vicini 2014, 143–149).

2. Healing

The example of deep brain stimulation shows how neuroscience is helping patients and could continue to offer approaches, technologies, and tools that could greatly benefit human health and social life. Health is a precious and limited good. Striving to promote healing is essential for individuals and the whole society, and it has been a constant commitment throughout the history of humankind and, particularly, within Christianity (Messer 2011; Kelly 2021).

Theological bioethics cares for the well-being of human beings and society, and it critically reflects on developments in neuroscience by engaging researchers, clinicians, civil servants, and citizens to examine current and proposed uses of technologies and tools in neuroscience aimed at improving the well-being of patients. Together with the development of scientific inquiry, both philosophical and theological discourses embrace healing as needed for personal and social flourishing.

Health and healing are comprehensive concepts and realities that encompass multiple dimensions. In the case of neuroscience, research that aims at deepening our study of brain and neurological functioning could have positive implications in healthcare contexts, by addressing the needs of patients, and by contributing to their well-being and quality of life. Hence, neuroscience could offer to healthcare professionals new possibilities to promote healing and health. In this case, by focusing on clinical applications resulting from studies and trials, as well as discoveries and tools, neuroscience will benefit one dimension of health, that is, healthcare practice.

Currently, it is not evident how developments in neuroscience could benefit two other synergic and essential approaches that contribute to promoting health locally and globally: public health and global health (Panicola and Barina 2018; Landrigan and Vicini 2021). The current global pandemic has made clear to the whole humankind how a comprehensive promotion of health demands a threefold engagement in promoting healthcare practice, public health, and global health. Bioethicists stress that this threefold engagement requires both personal and social commitments that allow the fostering of health at the institutional, structural, and systemic levels. In other words, it is not sufficient to focus on moral agents and their actions. A more comprehensive process requiring institutional, structural, and systemic transformations is needed locally, nationally, and globally.

3. Therapy and Enhancement

How we define health and healing has implications even as we consider applying the findings of research in neuroscience. In particular, one could wonder whether deep brain stimulators could be used not only in the case of intractable neuropathologies, but also in their absence in order to enhance specific human capabilities. At the same time, to expand our reflection, one could discuss what concerns cognitive enhancement, mood enhancers, and possible ways to influence moral enhancement (Messer 2017, 148–149; Reichlin 2019).

Applications of deep brain stimulators, as well as transcranial stimulators have been proposed not only to address the symptoms caused by neuropathologies, but also as a possible way to enhance the performance of individuals or influence their mood. For some authors, neuroenhancements are integral to their understanding of health and well-being.

In the ethical literature, both in philosophical and theological approaches, one finds examples of debates regarding the distinction between therapy and enhancement. The whole spectrum of positions can be identified. In recent years, these reflections were focused on genetic research and their applications (Gordijn, Chadwick *et al.* 2009). However, these positions and approaches could be applied to discussing issues regarding neuroscience.

On the one hand, for some authors the distinction between therapy and enhancement is quite clear. By relying on this distinction, they separate therapies from the pursuit of enhancement. While therapies are integral to the promotion of health, enhancements are considered unethical both for individuals and for the whole society because they introduce further dimensions of inequity in the social fabric (Habermas 2001; Sandel 2007). In this case, the ethical reasoning is attentive to considering the social implications of enhancing specific capabilities in human beings by challenging fairness and equality, and by avoiding fostering social inequities (Peterson-Iyer 2004, 170–210), as well as transhuman or posthuman seductions (Waters 2006; Cole-Turner 2011; Waters 2014).

On the other hand, one can find strong advocates of the pursuit of unbridled enhancement to express the human and social ability of fostering individual improvements even without considering the personal and social implications of pursuing these enhancements (Savulescu and Bostrom 2009; Clarke, Savulescu *et al.* 2016). The principle of respect of one's autonomy and individual freedom frame and support these positions (Buchanan 2011), even when social and political issues are addressed (Dubljević 2019).

In between these two opposite approaches one can also identify a middle position that is more cautious in assuming that it is always possible and very evident to distinguish between therapy and enhancement in each case. At the beginning of this paper, Sarah's story could be a clear example.

In her case, is deep brain stimulation a therapy or enhancement or both or neither? One could argue that it is not easy to separate between the clear ways in which her challenging and debilitating symptoms are improved by the stimulator and how her overall quality of life and well-being benefit from it at the point of being enhanced. At the same time, one cannot exclude that the brain simulation is also enhancing her abilities by giving her an extra edge compared to others who need to deal with their mood swings, as well as behavioral and personality aspects, without the stabilizing effects of a brain stimulator. Ethically, discernment is needed (Vicini 2012, 165–166).

Discernment implies considering the elements that characterize specific situations (i.e., the context and the circumstances), the means that are used, the moral agents involved, the intentions at play, what are the values and goods at stake, what will be the expected consequences and outcomes, and how one's experience, wisdom, and critical analysis illuminate each specific discernment process. Both examining and balancing these multiple factors characterize practical reasoning and they are informed by prudential judgment (Messer 2017, 167). Likely, one can expect that there will be diverse answers in different situations. This diversity, however, does not imply unbound moral relativism, as if anything goes and nothing matters. On the contrary, moral certitude is provisional, situated, limited, and it requires a continued reassessment and critique. Validation and confirmation are necessary, but they might also be provisional and open for possible revisions.

While this approach focuses on the moral agents and on their decision-making processes, other authors invite us to reflect critically on our understanding of enhancement. What do we consider as enhancement? In Sarah's case, should, we consider the stimulation and its benefits as an example of enhancement of her abilities? As Messer writes, "a good measure of epistemological humility is needed when we make claims about human flourishing and health" (Messer 2017, 161). However, what does epistemological humility entail in discussing the distinction between therapy and enhancement?

The American theologian James Keenan, S.J., for example, articulates his reflection on genetic enhancement by offering a critical reading of the

Christian tradition which focuses on the concept of perfection as integral to the Gospel and largely present in the history of Christianity, particularly in scholastic thought and in ascetic theology (Keenan 1999; Keenan 2007). He rightly distinguishes between, on the one hand, perfectionism (Kotva 1996, 16–47), as a self-centered and problematic striving that leads to "extreme individualism; an unhealthy preoccupation with perfect actions; a failure to consider the present; and extrinsic sources for determining one's perfection" (Keenan 1999, 110), and, on the other hand, the pursuit and gift of Christian perfection as a virtuous, integral dimension of discipleship. As Keenan writes, „The problem lies not with the question of *whether* we should pursue perfection, but rather *what* perfection we are pursuing" (Keenan 1999, 104)[2].

Following Keenan's invitation to consider the contributions of the Catholic tradition can also lead us to examine how both Catholic social thought and teaching expand our theological framework by inviting us to discuss individual behaviors, choices, and actions as well as social institution, structures, and systems in light of an understanding of social justice. This approach critically examines individual and social inequities and aims at fostering flourishing, solidarity, and the promotion of the common good with a preferential option for those who are excluded, marginalized, and discriminated (Hollenbach 2002; Cahill 2004; Cahill 2005, 235–239; Lysaught, McCarthy *et al.* 2018; Martins 2020).

4. Further Contributions

Further contributions should be mentioned. They show diverse directions of practical reasoning engaging neuroscience. For example, some authors are interested in exploring how neuropsychology (Jeeves and Ludwig 2018, 99–124) can contribute to our understanding of Christian spirituality, by stressing that spirituality can depend on our biological constitution (Jeeves and Ludwig 2018, 125–150). Other scholars prefer to articulate a neurotheology by suggesting that neuroscience will illuminate

[2] Italic in the original text.

our understanding of religion by unveiling its neuroscientific underpinnings (d'Aquili and Newberg 1999; Newberg and Waldman 2007; Geertz 2009; Goldberg 2009; Newberg and Waldman 2009; Newberg 2010; Joubert 2014; Kyriacou 2018; Manalili 2018; Newberg 2018; Newberg and Halpern 2018; Klemm 2019).

4.1 Protestant theological voices

Among Protestant theologians, for the American Alan Weissenbacher, neuroliteracy is needed, but neuroscience, and what it claims to offer us, "must be viewed with a critical eye" (Weissenbacher 2015, 48). As others suggested (Müller 2015), this careful approach demands to "discern the limits within which science works, and the limits to the capacity of neuroscience to explain just what makes us human in relationship to God" (Weissenbacher 2015, 48). By focusing on research in neuroscience, and how it aims at identifying what corresponds, in our brain, to our behaviors and actions, Weissenbacher stresses how

> the fact that one discovers neural correlates of human experience should be expected. The threat of reductionism comes when one uses these discoveries to advance a reductive and one-dimensional view of human nature, explaining away an aspect of human experience because it has a physical substrate (Weissenbacher 2015, 48).

Moreover, for Weissenbacher, researchers should be aware that "phenomena as complex as mental processes" (Weissenbacher 2015, 48) require avoiding unhelpful simplifications. In particular, the current technologies that we use to study brain functioning and brain pathologies are simply "processed and averaged representations of activity" (Weissenbacher 2015, 48) and "FMRI scans do not literally read thoughts, they measure brain oxygen levels" (Weissenbacher 2015, 48). He continues,

> additionally, a point of brain activation may not indicate a seat of a particular behavior, as it could rather be distributed over several brain activity regions. Realistic conclusions regarding brain function and mental pro-

> cesses should incorporate the plastic nature of the brain as a non-linear, dynamic system with parallel processes and redundancy with brain regions that overlap and are multifunctional (Weissenbacher 2015, 48).

In his 2017 volume *Theological Neuroethics: Christian Ethics Meets the Science of the Human Brain*, the British theologian Neil Messer articulates his engagement with neuroscience by discussing: neuroscientific accounts of religion; the neuroscience of morality; questions about free will and responsibility, raised by neuroscience since the 1980s, and that engage our understanding of God's providence, sin, and salvation while in dialogue with Augustine (354–430) and Karl Barth (1886–1968); and the ethics of modifying brain functions technologically, either for therapeutic or enhancement purposes.

In engaging the field of neuroethics (Giordano and Gordijn 2010; Racine 2010; Racine and Aspler 2017), Messer suggests that we should consider, on the one hand, the *ethics of neuroscience*, that is the ethical conduct of neuroscientific research and the use of its findings, but also their ethical implications. On the other end stands the *neuroscience of ethics*, that is, the scientific study of moral reasoning, judgement, and action (Messer 2017, 3; Messer 2021, 351–352).

Regarding neuroscience and religion, in articulating his interdisciplinary approach Messer's threefold goal is, first, to engage the cognitive science of religion,

> whose aim is to identify cognitive mechanisms that give rise to religious or supernatural beliefs; second, evolutionary theories about the origins of religious beliefs and practices; and third, neuroscientific attempts to identify brain activity and mechanisms associated with religious beliefs, practices and experiences (Messer 2017, 15).

This approach leads him to affirm that

> the scientific study of the brain does not give good reasons to dismiss theological claims or reasoning, and that theological engagements with neuroscience are better framed in a way that gives priority to theological sources

> and perspectives than by handing over control of the account to scientific perspectives (Messer 2017, 15–16).

In examining possible typologies regarding how sciences can study what concerns religion, Messer reclaims an approach that prioritizes theological reasoning in articulating a critical assessment of religious discourse and experience. However, such a priority does not imply dismissing any significant contribution offered by the three approaches that he discusses: cognitive sciences, evolutionary theories, and neuroscience. In his words, he aims at testing

> the capacity of a Christian theology firmly rooted in its own distinctive sources and methods to appropriate insights critically from evolutionary, cognitive and neuroscientific research, to incorporate them in a theologically-formed vision, and so to articulate a distinctive and compelling theological neuroethics (Messer 2017, 35–36).

4.2 Roman Catholic theological voices

While some authors prefer to reflect on brain science, the mind, human nature (Murphy 2002; Dodson 2017), and the soul (Hess and Allen 2008, 165–167; Gay 2009; Clarke 2015, 149–215; Crisp, Porter *et al.* 2016; Lombard 2017), in the Catholic context Christopher Vogt examines how recent scholarship that engages neuroscience focuses on virtues and "illuminates the dynamics of personal formation and social transformation" (Vogt 2016, 182).

While Vogt is suspicious of many assertions made by neuroscientists regarding the explanatory power of their studies on human decision making and moral life, as well as any direct and uncritical application of findings from animal studies to human beings, he considers more positively the possible contributions offered by neuroscience to study human emotions and how they contribute to shaping moral agency (Vogt 2016, 186–187). In doing so, he joins recent philosophical and theological debates that stress the rationality of human emotions and how our emotional being contributes to moral reasoning and influences

moral action (Nussbaum 2001; Cates 2009; Cochran 2015; Vogt 2016, 186; Jaycox 2021).

Strengthening Vogt's critique, one can wonder which understanding of morality appears to inform research protocols, experiments, and their results. In too many instances, researchers do not seem to appreciate the richness of the ethical tradition, with its philosophical and theological sources. In very reductive ways, for many neuroscientists two approaches sum up the whole ethical tradition: first, the trolley problem; second, the categorical imperative proposed by Immanuel Kant (1724–1804): "Act only according to that maxim by which you can at the same time will that it should become a universal law" and "So act as to treat humanity, whether in your own person or in another, always as an end and never as only a means" (Kant 2015).

The trolley case was first proposed by philosopher Philippa Foot (1910–2010) in 1967 (Foot 1967; Foot 1978a; Foot 1978b, 19–32) and then adapted by philosopher Judith Jarvis Thomson (1929–2020) a few years later (Thomson 1976; Thomson 1985). With its two hypothetical scenarios the trolley problem is considered able to exemplify, on the one hand, the struggle of the moral agent's decision-making progress in challenging situations in which any outcome is ethically problematic and troubling. On the other hand, what is also supposed to be assessed is how the proximity or distance of one moral agent with other moral agents might influence one's decision regarding how to act.

However, to reduce ethical decision making to choosing among the two proposed choices that frame the trolley case, without considering anything else and without allowing for proposing diverse ways to handle those hypothetical situations, seems to impoverish what it means to be a moral agent capable of making choices, what are the ethical resources that allow us to address critically challenging situations with undesirable outcomes, how we discuss one's responsibility, and how we examine critically social, cultural, and religious contexts in which one's decisions occur (Fried 2012; Hildt 2015, 65–68; Kamm, *et al.*, with *et al.* 2016; Messer 2017, 45).

To assess moral decision making and agency requires a more nuanced, diversified, and comprehensive assessment that examines how the ability of the moral agent of making choices, and of assessing the consequences of de-

cisions and actions, depends on considering one's intentions, means, and circumstances as well as the multiple elements that characterize the context in which decisions and actions occur (Messer 2016, 55–58).

Moreover, while the Kantian categorical imperative represents a very significant achievement in the history of ethical reasoning, by considering it the exclusive element that guides one's moral decision-making process and action impoverishes the importance that experience, discernment, and tradition-with their contributions-have in forming the personal and social ethos and in shaping agency. Within the ethical tradition, principles and norms, virtues and vices, rights and duties play very important roles in describing and defining moral agents, their practical reasoning abilities, their decision-making capabilities, their critical assessment of decisions and actions as well as of contexts, structures, systems, and dynamics.

Concretely, while one welcomes collaborative projects in research in neuroscience that engage both ethicists and neuroscientists, interdisciplinary interactions and collaborations seem to be needed because expertise and disciplinary mastery of both the ethical tradition and neuroscience neither could be presumed nor required. Such an interdisciplinary interaction could allow neuroscientists to study moral decision making and agency in ways that are nourished and inspired by the richness of the ethical tradition, by avoiding reductive and oversimplifying approaches that are exclusively centered on a very narrow understanding of ethics by reducing it to the trolley problem and to the Kantian categorical imperative.

In this vein, Vogt argues that

> the excesses and lack of philosophical sophistication of many works on morality and neuroscience may have discouraged some theologians from engaging that field, but several recent works show the promise of neuroscience as a dialogue partner for theology, and for virtue ethics in particular (Vogt 2016, 186).

Hence, Vogt joins the recent renewed interest in virtue ethics, both among philosophers (Russell 2013) and theologians (Kotva 1996; Curran and Fullam 2011; Porter 2012; Salzman and Lawler 2018; Daly 2021a).

Furthermore, for Vogt "research in neuroscience can be quite helpful in terms of supporting the empirical claims about moral formation as well as how human beings engage in moral reasoning and the role emotions play in those processes" (Vogt 2016, 186). He continues,

> multiple studies of individuals who suffered damage only to areas of the brain that are strongly associated with emotion without any damage to non-affective neural areas exhibited significant impairments in social judgment, practical reasoning, and inter-personal relationships (Vogt 2016, 186–187).

While researchers are confirming these assumptions (Bowers and Yehuda 2016; Bowren, Croft *et al*. 2018; Cardinale, Reber *et al*. 2021) and stimulate further inquiry, Vogt also adds that "recent studies in neuroscience and the behavioral sciences have emphasized the importance of exemplars" (Peterson 2012; Vogt 2016, 188) and how these exemplars contribute to inspire moral agents and form them. Moreover, he comments on how, for some scholars (Burns 2012; Brown and Reimer 2013), narratives can inform our behavior and agency. In particular, "by implicitly simulating in one's own sensorimotor systems the actions being narrated" (Brown and Reimer 2013, 843; Vogt 2016, 189), what neuroscientists appear to identify confirms the importance that stories and narratives have in contributing to moral agency, as it is stressed by scholars interested in exploring the importance of literature studies in ethics (Hauerwas and Jones 1989; Rowe and Horner 2010).

5. Conclusion

Some authors turn to neuroscience with the purpose of applying results of its studies in specific fields, from higher education (Baker and Leonard 2017) to national security (Giordano 2015). These examples further confirm the need for theological bioethics to join collaborative efforts aimed at examining the research agenda of neuroscientists and to join their efforts by contributing with a thick series of ethical concerns and resources that

critically strive to promote virtuous behaviors in the moral agents and virtuous dynamics in society.

Moreover, theological bioethics can engage in interdisciplinary interactions in neuroscience because theologians are animated by, first, an *anthropological vision* that makes explicit the need to respect human dignity in its vulnerability and fragility, as an expression of being created in the image of God and being capable of virtuous moral agency. Second, this vision of the person is further enriched by a *vision of science* that rejects any abusive and elitist approach to scientific investigation, research, and implementation, and that is animated by a search for knowing that is critically informed. Finally, a *vision of society* shaped by just relations, attentive to unmask unjust power dynamics, and striving to foster virtuous structures (Daly 2021b) will greatly contribute to situate developments in neuroscience in a transformed social fabric that longs to reducing and eliminating social inequities.

Bibliography

Baker, D. L., Leonard, B. Neuroethics in Higher Education Policy, New York 2017.

Belluck, P., A 'Pacemaker for the Brain? No Treatment Helped Her Depression-until This, "New York Times", October 4, 2021.

Bowers, M. E. and R. Yehuda, Intergenerational Transmission of Stress in Humans, "Neuropsychopharmacology", 2016, 41, 1, 232–244.

Bowren, M. D. *et al.*, Choosing Spouses and Houses. Impaired Congruence between Preference and Choice Following Damage to the Ventromedial Prefrontal Cortex, "Neuropsychology", 2018, 32, 3, 280–303.

Brown, W. S., Reimer, K. S., Embodied Cognition, Character Formation, and Virtue, "Zygon", 2013, 48, 3, 832–845.

Buchanan, A. E., Beyond Humanity? The Ethics of Biomedical Enhancement, Oxford/New York 2011.

Burns, C. P. E., Hardwired for Drama? Theological Speculations on Cognitive Science, Empathy, and Moral Exemplarity, New York 2012.

Cahill, L. S., Bioethics and the Common Good, Milwaukee, WI 2004.

Cahill, L. S., Theological Bioethics. Participation, Justice, and Change, Washington, D.C. 2005.

Cardinale, E. M. *et al.*, Bilateral Amygdala Damage Linked to Impaired Ability to Predict Others' Fear but Preserved Moral Judgements About Causing Others Fear, "Proceedings of Royal Society B: Biological Sciences", 2021, 288, 20202651.

Cates, D. F., Aquinas on the Emotions. A Religious-Ethical Inquiry, Washington, D.C. 2009.

Clarke, P. G. H., All in the Mind? Challenges of Neuroscience to Faith and Ethics, Oxford 2015.

Clarke, S. *et al.* (eds), The Ethics of Human Enhancement. Understanding the Debate, Oxford 2016.

Cochran, E. A., The Moral Significance of Religious Affections. A Reformed Perspective on Emotions and Moral Formation, "Studies in Christian Ethics", 2015, 28, 2, 150–162.

Cole-Turner, R., (ed.), Transhumanism and Transcendence. Christian Hope in an Age of Technological Enhancement, Washington, D.C. 2011.

Crisp, T. M., Porter, S. L., Ten Elshof, G. (eds), Neuroscience and the Soul. The Human Person in Philosophy, Science, and Theology, Grand Rapids, MI 2016.

Curran, C. E. Fullam, L. (eds), Virtue (=Readings in Moral Theology no. 16), Mahwah, NJ 2011.

d'Aquili, E. G., Newberg, A. B., The Mystical Mind. Probing the Biology of Religious Experience, Minneapolis, MN 1999.

Daly, D., Virtue Ethics, in: Winright, T. L. (ed.), T&T Clark Handbook of Christian Ethics, London/New York 2021a, 59–70.

Daly, D. J., The Structures of Virtue and Vice, Washington, D.C. 2021b.

Dodson, G. F., Free Will, Neuroethics, Psychology and Theology, Wilmington, DE 2017.

Dubljević, V., Neuroethics, Justice and Autonomy. Public Reason in the Cognitive Enhancement Debate, New York 2019.

Foot, P., The Problem of Abortion and the Doctrine of the Double Effect, "Oxford Review", 1967, 5, 5–15.

Foot, P., The Problem of Abortion and Negative and Positive Duty. A Reply to James Leroy Smith, "Journal of Medical Philosophy", 1978a, 3, 3, 253–255.

Foot, P., Virtues and Vices, and Other Essays in Moral Philosophy, Oxford 1978b.

Fried, B. H., What Does Matter? The Case for Killing the Trolley Problem (or Letting It Die), "Philosophical Quarterly", 2012, 62, 248, 505–529.

Gay, V. P., (ed.), Neuroscience and Religion. Brain, Mind, Self, and Soul, Lanham, MD 2009.

Geertz, A. W., When Cognitive Scientists Become Religious, Science Is in Trouble. On Neurotheology from a Philosophy of Science Perspective, "Religion", 2009, 39, 4, 319–324.

Giordano, J. J., (ed.), Neurotechnology in National Security and Defense. Practical Considerations, Neuroethical Concerns. Advances in Neurotechnology, Boca Raton, FL 2015.

Giordano, J. J., Gordijn, B. (ed.), Scientific and Philosophical Perspectives in Neuroethics, Cambridge, UK/New York 2010.

Goldberg, D. W., D'aquili and Newberg's Neurotheology. A Hermeneutical Problem with Their Neurological Solution, "Religion", 2009, 39, 4, 325–330.

Gordijn, B., Chadwick, R. (ed.), Medical Enhancement and Posthumanity (=The International Library of Ethics, Law and Technology), Dordrecht 2009.

Habermas, J., Die Zukunft Der Menschlichen Natur. Auf Dem Weg Zu Einer Liberalen Eugenik?, Frankfurt am Main 2001.

Hauerwas, S., Jones, L. G. (eds), Why Narrative? Readings in Narrative Theology, Grand Rapids, MI 1989.

Hess, P. M. J., Allen, P. L., Catholicism and Science, Westport, CT 2008.

Hildt, E., Brain, Morals, and Ethics. What Is the Connection?, "Concilium: Theology, Anthropology, and Neuroscience", 2015, 4, 63–71.

Hollenbach, D., The Common Good and Christian Ethics, Cambridge, UK/New York 2002.

Jaycox, M. P., Emotions and Christian Ethics, in: Winright, T. L. (ed.), T&T Clark Handbook of Christian Ethics, London/New York 2021, 91–101.

Jeeves, M. A., Ludwig, T. E., Psychological Science and Christian Faith. Insights and Enrichments from Constructive Dialogue, West Conshohocken, PA 2018.

Joubert, C., Christians, the Brain, and Person. Conceptual Confusion, Unintelligibility, and Implications, "Answers Research Journal", 2014, 7, 189–204.

Kamm, F. M. *et al.*, The Trolley Problem Mysteries, New York 2016.

Kant, I., Critique of Practical Reason, Cambridge, UK 2015.

Keenan, J. F., 'Whose Perfection Is It Anyway?'. A Virtuous Consideration of Enhancement, "Christian Bioethics", 1999, 5, 2, 104–120.

Keenan, J. F., Perfecting Ourselves. On Christian Tradition and Enhancement, "Southern Medical Journal", 2007, 100, 1, 96–97.

Kelly, C. M., Medicine, Bioethics, and Health Care, in: Winright, T. L. (ed.), T&T Clark Handbook of Christian Ethics, London/New York 2021, 299–307.

Klemm, W. R., Whither Neurotheology?, "Religions", 2019, 10, 11.

Kotva, J. J., The Christian Case for Virtue Ethics, Washington, D.C. 1996.

Kyriacou, D., Are We Wired for Spirituality? An Investigation into the Claims of Neurotheology, "HTS Teologiese Studies/Theological Studies", 2018, 74, 3.

Landrigan, P. J., Vicini, A. (eds), Ethics of Global Public Health. Climate Change, Pollution, and the Health of the Poor, Eugene, OR 2021.

Lombard, J., The Mind of God. Neuroscience, Faith, and a Search for the Soul, New York 2017.

Lysaught, M. T., McCarthy, M. P., Cahill, L. S. (eds), Catholic Bioethics and Social Justice. The Praxis of Us Health Care in a Globalized World, Collegeville, MN 2018.

Manalili, M. M. C., On Neurotheology? Why Engage Empirical Studies on Theological Concepts, "Lumen et Vita", 2018, 9, 1.

Martins, A. A., The Cry of the Poor. Liberation Ethics and Justice in Health Care, Lanham, MD 2020.

Messer, N., Toward a Theological Understanding of Health and Disease, "Journal of the Society of Christian Ethics", 2011, 31, 1, 161–178.

Messer, N., Cognitive Science, Moral Reasoning, and the Theological Suspicion of Ethics, "Journal of the Society of Christian Ethics", 2016, 36, 1, 51–68.

Messer, N., Theological Neuroethics. Christian Ethics Meets the Science of the Human Brain, London 2017.

Messer, N., Neuroethics, in: Winright, T. L. (ed.), T&T Clark Handbook of Christian Ethics, London/New York 2021, 351–358.

Müller, K., The Dulcet Tones of a New Doctrine of Salvation? On the Achievements and Limnits of Neurosciences, "Concilium: Theology, Anthropology, and Neuroscience", 2015, 4, 35–45.

Murphy, N., Neuroscience and Human Nature. A Christian Perspective, in: Peters, T., Iqbal, M., Haq, S. N. (eds), God, Life, and the Cosmos. Christian and Islamic Perspectives, Aldershot/Burlington, VT 2002, 357–381.

Newberg, A. B., Principles of Neurotheology, Farnham, Surrey, England/Burlington, VT 2010.

Newberg, A. B., Neurotheology. How Science Can Enlighten Us About Spirituality, New York 2018.

Newberg, A. B., Halpern, D., The Rabbi's Brain. An Introduction to Jewish Neurotheology, Nashville, TN 2018.

Newberg, A. B., Waldman, M. R., Born to Believe. God, Science, and the Origin of Ordinary and Extraordinary Beliefs, New York 2007.

Newberg, A. B., Waldman, M. R., How God Changes Your Brain. Breakthrough Findings from a Leading Neuroscientist, New York 2009.

Nussbaum, M. C., Upheavals of Thought. The Intelligence of Emotions, Cambridge, UK/New York 2001.

Panicola, M., Barina, R., Catholic Health Care and Population Health. Insights from Catholic Social Thought, in: Lysaught, M. T., McCarthy, M. P. (eds), Catholic Bioethics and Social Justice. The Praxis of Us Health Care in a Globalized World, Collegeville, MN 2018, 283–296.

Peterson, G. R., Virtue, Science, and Exemplarity, in: Van Slyke, J. A. *et al.* (eds), Theology and the Science of Moral Action. Virtue Ethics, Exemplarity, and Cognitive Neuroscience, New York 2012, 27–46.

Peterson-Iyer, K., Designer Children. Reconciling Genetic Technology, Feminism, and Christian Faith, Cleveland, OH 2004.

Porter, J., Virtue Ethics, in: Gill, R. (ed.), The Cambridge Companion to Christian Ethics, Cambridge, UK/New York 2012, 87–102.

Racine, E., Pragmatic Neuroethics. Improving Treatment and Understanding of the Mind-Brain, Cambridge, MA 2010.

Racine, E. Aspler, J. (eds), Debates About Neuroethics. Perspectives on Its Development, Focus, and Future (=Advances in Neuroethics), Cham, Switzerland 2017.

Reichlin, M., Neuroscienze, in: Benanti, P. *et al.* (eds), Teologia Morale, Cinisello Balsamo, MI 2019, 649–656.

Rowe, A., Horner, A. (eds), Iris Murdoch and Morality, Basingstoke, Hampshire/New York 2010.

Russell, D. C., (ed.), The Cambridge Companion to Virtue Ethics. Cambridge, UK/New York 2013.

Salzman, T. A., Lawler, M. G., Virtue and Theological Ethics. Toward a Renewed Ethical Method, Maryknoll, NY 2018.

Sandel, M. J., The Case against Perfection. Ethics in the Age of Genetic Engineering, Cambridge, MA 2007.

Savulescu, J., Bostrom, N., Human Enhancement, Oxford/New York 2009.

Scangos, K. W. *et al.*, Closed-Loop Neuromodulation in an Individual with Treatment-Resistant Depression, "Nature Medicine", 2021, 27, 1696–1700.

Thomson, J. J., Killing, Letting Die, and the Trolley Problem, "The Monist", 1976, 59, 2, 204–217.

Thomson, J. J., The Trolley Problem, "Yale Law Journal", 1985, 94, 6, 1395–1415.

Vicini, A., Ethical Challenges of Human Genetics Today. From the Lab, through the Clinic, to the Pews, "Studia Moralia", 2012, 50, 1, 145–174.

Vicini, A., Le Neuroscienze e la Bioetica, "La Civiltà Cattolica", 2014, 163/II, 3932, 143–158.

Vogt, C. P., Virtue. Personal Formation and Social Transformation, "Theological Studies", 2016, 77, 1, 181–196.

Waters, B., From Human to Posthuman. Christian Theology and Technology in a Postmodern World, Aldershot, Hants, England/Burlington, VT 2006.

Waters, B., Christian Moral Theology in the Emerging Technoculture. From Posthuman Back to Human, Farnham, Surrey/Burlington, VT 2014.

Weissenbacher, A. C., Ten Principles for Interpreting Neuroscientific Pronouncements Regarding Human Nature, "Dialog: A Journal of Theology", 2015, 54, 1, 40–50.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 45–63
DOI: 10.24989/BCE.neuro.3

Neurosciences and Theological Ethics: Remarks to an Interdisciplinary Discourse

Neurowissenschaften und theologische Ethik: Anmerkungen zu einem interdisziplinären Diskurs

Angelika Walser, Salzburg

Abstract (Deutsch)

Mein Beitrag versucht einen interdisziplinären Dialog zwischen den Neurowissenschaften und der theologischen Ethik. Ich beginne mit Überlegungen zur Erzählung von Phineas Gage, dessen Geschichte manchmal als spektakulärer Beweis für die Nichtexistenz des freien Willens dient. Ich hinterfrage die Erzählung von Gage, indem ich sie mit einer alltäglichen Geschichte meiner heranwachsenden Tochter vergleiche und beide Geschichten verwende, um die Aufgaben aufzuzeigen, die die Neurowissenschaften der Ethik erteilt. Es liegt auf der Hand, dass es eine neuronale Grundlage für moralische Entscheidungen gibt, die die theologische Ethik nicht ignorieren kann, wenn sie über moralische Entwicklung und moralisches Urteilsvermögen spricht. Neuronale Prozesse sind zwar eine notwendige, aber nicht hinreichende Komponente für die Entscheidungsfindung bzw. das soziale und moralische Verhalten. Unter Bezugnahme auf einen Schlüsseltext der selbstreflexiven Neurowissenschaft (Memorandum) weise ich auf drei Aspekte hin, die ich für einen erfolgreichen interdisziplinären Diskurs aus der Perspektive der theologischen Ethik für wichtig halte: erstens, die Notwendigkeit der Einbettung der Neurowissenschaften in einen gesellschaftlichen Kontext; zweitens die notwendige Kritik an Kategorienfehlern und das Bewusstsein für unterschiedliche Perspektiven, Vorgehensweisen und Grenzen beider Disziplinen; drittens, die Bedeutung einer nicht-dualistischen und relationalen Anthropologie für den freien Willen.

Abstract (English)
My contribution attempts an interdisciplinary dialogue between neurosciences and theological ethics. I start with reflections on the narrative of Phineas Gage whose story functions sometimes as a spectacular proof of the non-existence of the free will. I question the narrative of Gage by comparing it to an everyday story of my adolescent daughter, using both stories to introduce the lessons that neurosciences teach ethics. It is obvious that there is a neural base for moral decision-making, which theological ethics cannot ignore when talking about moral development and moral judgment. However, neuronal processes are a necessary but not sufficient component for decision-making resp. social and moral behaviour. Referring to a key text of self-reflexive neuroscience (*Memorandum*), I point to three aspects that I consider important for a successful interdisciplinary discourse from the perspective of theological ethics. First, the necessity of embedding neurosciences in a social context. Second, the necessary critique of category mistakes and the awareness of different perspectives, methodologies, and limits of both disciplines. Third, the importance of a non-dualist and relational anthropology for free will.

Keywords (Deutsch)
Neurowissenschaften; interdisziplinärer Dialog; theologische Ethik; moralische Verantwortung; freier Wille;

Keywords (English)
neurosciences; interdisciplinary dialogue; theological ethics; moral responsibility; free will;

1. Introduction: Phineas Gage and My Daughter

One of the current grievances of humanity (and theology) is the supposed "abolition of the freedom of will", caused by neurosciences. Usually, the famous story of Phineas Gage is used to prove this thesis. The story of the railway site manager, whose skull was pierced in an accident with an iron bar in 1848, is common knowledge: Although the bar had destroyed

a large part of his left brain (the orbitofrontal cortex), Gage survived the accident. He recovered soon from his physical injuries and his cognitive capabilities seemed to be intact. However, after a while his colleagues and friends discovered a change in his personality: Before his accident, Gage had been a quiet and reflective man. Afterwards, he showed aggressive, impulsive, and unreliable behaviour. According to the common narrative, he lost several jobs and died in social isolation and material poverty. His "horror story", sad and fascinating at the same time, seems to underline that from a neuroscientific view both reasoning and social capacities seem to be a question of brain functions only. On an even more fundamental level, philosophical and theological assumptions of the existence of free will as a base for human morality had supposedly turned out to be pure illusion. At least, this is what 10 male and 1 female leading representatives of neurosciences in the German-speaking world claimed in their *Manifesto about Present and Future of Brain Research*. They promised enormous progress in fighting diseases like Alzheimer and Parkinson, schizophrenia and depression, mental abnormalities and maldevelopment. Finally, they predicted, "that mind, consciousness, feelings, acts of will and freedom of action are regarded as natural processes without contradiction, because they are based on biological processes." (Monyer *et al.* 2004, 36). In fact, readers of the *Manifesto* had to gain the impression that neurosciences had not only discovered various neural mechanisms, but the key to the explanation of human being in its entirety.

Ten years later, 2014, there is a certain disillusionment: In remembrance of the manifesto of 2004, there was a new manifesto, the so-called *Memorandum of Reflexive (i.e., Thoughtful) Neuroscience*. The list of signatories, which now consisted exclusively of men, included not only neuroscientists, psychologists, and psychiatrists, but also a few representatives of philosophy and ethics. The text itself took a sober stock of brain research and criticized many expectations of the 2004 *Manifesto* as completely exaggerated:

> From our point of view, however, today's balance sheet is rather disappointing. There is no rapprochement with the goals set in sight. The reasons for this go far beyond organizational-technical difficulties and lie on the one

> hand in weaknesses in the field of neuroscience theory, and on the other hand in naturalistic presuppositions and concepts that are not well thought out enough, which make desirable bridges to psychology, philosophy and cultural studies difficult in the long term. (Tretter *et al.* 2014, 1)

In spite of clear and undeniable progress in the field of neurosciences, many representatives have become thoughtful, aware of limitations, and humble:

> Ultimately, the reduction of humans and all their intellectual and cultural achievements to their brains is completely inadequate as a 'new conception of human'. This one-sided grid cannot grasp human being as a subject and person in its complexity. It is always the whole person who perceives, considers, decides, remembers, etc., and not a neuron or a cluster of molecules. (Tretter *et al.* 2014, 2)

At the end of the *Memorandum*, there is a request for interdisciplinary dialogue, especially with philosophy and ethics. In the following contribution I would like to accept this invitation, but as a representative of *theological* ethics. I am fully aware that neither the *Manifesto* nor the *Memorandum* have ever mentioned theology. However, theological ethics is still alive and aware of both, the discourse in neurosciences and philosophy, and there might be a special contribution from the perspective of theology as well.

Before I start to outline the lessons, which theological ethics could learn from neurosciences and, vice versa, some theological-ethical remarks about the relevance and impact of neurosciences, I would like to tell a little story: It is not as famous and spectacular as the one about Phineas Gage. However, from a neuroscientific point of view it is closer to Phineas Gage's story than one might initially expect. It is a story about my 14-year-old daughter: On a sunny day, her class teacher gave me a call to inform me that she had left the area of the school, together with her girlfriend, during a break. She had not returned to class and the teacher wondered if she had gone home. In fact, she did not arrive home and in the course of the day I became worried. At the end, the whole school was searching for the two girls, the principal called me and we were highly alarmed. At six

o'clock in the evening, I went to see a neighbour who, in former times, used to be her babysitter. To my surprise, I met my daughter at the front of her house, relaxed, smiling, and talking to the neighbour, her hands full of bags. I asked her: "Where the hell have you been all the time?" and felt like doing what good mothers should not do. She replied: "Hey Mom, come on. We felt like going shopping today. Just relax, we had a good time!" When I told her about the trouble at school, about our worries and fears, she was shocked: "Sorry, Mom, I haven't thought of all these things. I simply forgot you. I am sorry. We just felt like having fun!" After recovering from the first shock, I searched for an explanation for the strange and ignorant behaviour of a girl who had been quite a socially intelligent child in former times. In fact, the phenomenon of memory gaps and unsocial resp. immoral behaviour of young adults is a well-known issue in literature. I stumbled across a book written by the German cognitive developmental psychologist Eveline Crone who gives a good overview of the latest research on the adolescent brain.

1.1 Lessons from Neurosciences: The Neuroanatomy of Morality

Crone summarizes the results of research in moral neuroscience: According to this research, children and adolescents at the age between 6–20 are quite similar to Phineas Gage (Crone 2016, 104–109). Although highly intelligent, they are often not able to predict the long-term consequences of their decisions and live in the moment, guided by their emotions. Crone cites Antonio Damasio's "somatic markers hypothesis" (Damasio 1997, 227–273) and explains anti-social behaviour, which is typical for *all* teenagers, with these missing somatic markers. Usually, these markers link a special situation with a special feeling and are an essential part of our capacity to meet decisions in a very short time and in an intuitive way. They are a kind of physical signal that gives us a sense of how to decide. Many international studies proved that the development of these markers is due to complex restructuring processes of the orbitofrontal cortex. Only at the age of 16 to 18, these markers start to develop and they take many years, up to the age of 25, to reach an adult level. Thus, the temporary disorders in emotion-based decision-making and sometimes even highly destructive

behaviour of teenagers result from changes in the brain. Crone dedicates one chapter of her book to what she calls the *social* brain and another to the *moral* brain. In the chapter about the *social* brain (Crone 2016, 133–146), she reports about tumours affecting the orbitofrontal cortex. They can leave the young patient's intellectual abilities intact while destroying the ability to build up friendships with others by antisocial and immoral behaviour like lying and cheating like Gage. Crone reports as well that it takes quite a long time in general until the neuronal basis allows children to empathize with others and put themselves in other's shoes.

In the chapter about the *moral* brain, Crone reports that Lawrence Kohlberg and Jean Piaget's concepts about the development of moral judgment towards moral autonomy receive neural confirmation (Crone 2016, 146–150). Solving moral dilemma situations, like referring to universal principles and ideals on a post-conventional level, requires a highly developed neural basis. However, it is important to mention that while especially Kohlberg's studies represent a milestone in the psychological study of morality, they

> considered moral reasoning to be a result of cognitive processes that may exist even in the absence of any kind of emotions. However, findings in evolutionary psychology […] and primatology […] suggested that emotions played a key part in the origins of human morality. (Pascual *et al.* 2013, 2)

Of course, Crone's description of the "moral brain" is due to a popular science presentation.

The neural circuits of brain regions implicated in morality overlap with those that regulate other behavioral processes, suggesting that there is probably no undiscovered neural substrate that uniquely supports moral cognition. The most plausible option is that the 'moral brain' does not exist *per se*: rather, moral processes require the engagement of specific structures both the 'emotional' and the 'cognitive' brains, […]. (Pascual *et al.* 2013, 5)

Research goes on: In well-known studies about moral dilemma situations, such as the trolley situation, Joshua Greene investigated which

brain areas are involved in moral decision-making (Greene 2016, 1–15). He could prove that the emotional areas were quite active while thinking about a personal dilemma, while reflection about abstract and impersonal moral problems activated cognitive areas. He could also prove on a neural basis what each ethicist knows from his/her own experience: The more you are emotionally involved in a moral problem the more difficult it is to come to a decision.

The exact role of emotions in the field of morality still requires further research. According to Avramova and Inbar, there are three claims: The first and least controversial claim,

> emotions follow from moral judgment, such that witnessing immorality can lead to negative emotions and witnessing moral virtue can lead to positive ones. According to the second claim, emotions amplify moral judgements, for instance, by making immoral acts seem even more immoral. Finally, on the last claim, emotions can actually moralize nonmoral behaviors – that is, they give nonmoral acts a moral status. (Avramova & Inbar 2013, 169)

While there is clear support for the first claim, there is only limited empirical evidence for the second and the third claim. Of course, it is not surprising that emotions are a consequence of (im)moral behaviour. Whether emotions are also a source or even a predictor of moral judgment, is still an open discussion. In any case, there seems to be enough empirical evidence to state that strong emotions, as for instance physical disgust, elicit negative moral judgment (Avramova & Inbar 2013, 173). Moral outrage seems to guide perception, moral decision-making, and moral behaviour. In sum, one can assume that emotions are much more than a mere consequence of experience. They probably guide our perception and prioritising of moral problems.

1.2 Reactions on the Part of (Theological) Ethics

It is evident that neuroscience has important implications for moral decision-making and moral judgment. To sum it up in the words of the aforementioned *Memorandum*:

"Without the brain everything is nothing" (Tretter *et al.* 2014, 1)! It is obvious that there is a neural base for moral-decision making which has to be taken into account when assessing moral maturity and moral responsibility of agents. Age and the stage of cognitive development are not only legal issues affecting criminal responsibility, but also issues of morality and ethics. One cannot expect the highest grades of autonomous decision-making and moral responsibility from people who lack a part of their neuronal capacity due to youth, accident, trauma, or cognitive impairment. "Brain dysfunction diminishes or undermines responsibility when it diminishes and undermines the physical and mental capacities necessary for responsibility" (Glannon 2011, 17). It would be a mistake, though, to conclude that a dysfunctional brain is decisive for the issue of moral responsibility, so that, e.g., a psychopath suffering from partial brain dysfunction cannot be hold responsible at all for what he is doing and that it simply takes some brain stimulation to make him become a morally better person in the sense of being less cruel and more sensitive towards others. Moral responsibility also entails conscious mental states and the awareness of being the originator of one's own decisions and actions. Studies show that even psychopaths do have a knowledge about their mental states and are not compelled to act (Glannon 2011, 19ff). On the other hand, one should not fall into the opposite trap and reduce moral capabilities to mere cognitive and conscious capacities. Studies conducted by Antonio Damasio have shown, that emotions, complex physical reactions which cannot be controlled consciously, are a necessary part of our thinking and our moral decision-making. According to theological ethicist Michael Rosenberger's summary of Antonio Damasio's research in his introduction to moral theology, titled *Frei zu leben!*, emotions represent condensed value experiences (Rosenberger 2018, 336–338). They are carriers of information and interpretation and are as cognitive as any other form of perception. They are complex stereotyped response patterns and a consequence

of feelings. According to Damasio, who distinguishes between "emotions" and "feelings", the latter are sensory patterns that signal pain and pleasure and become imaginations (Damasio 2006, 73). Feelings are representations of bodily states in the brain and induce emotions. Rational thinking as a precondition for morality depends on these emotions, which enable human beings to focus on central values. Emotions are inseparably linked with the idea of punishment and award, pleasure and pain, good and evil (Damasio 2006, 72). Thus, moral responsibility presumes an interaction of both, conscious mental states and unconscious bodily processes including emotions. It is not an individualist project, but closely linked to society and culture, which consider certain moral values to be of central importance and shape these values by education.

In terms of moral pedagogy, from a neuroscientific perspective, Martha Nussbaum is right to claim that democratic societies of the 21st century do not need only justice and respect for one another, grounded in the Kantian concept of respect for human dignity. They also need passionate love, taught and cultivated in public places, like universities and schools, museums and parks (Nussbaum 2016, 569ff). Emotions and passions, traditionally assigned to the feminine or right brain, are not disturbing elements, which trouble the calm and clear rationalist left brain mind of male philosophers. They are an integral part of practical reasoning. Morality is not a matter of distant reflection, a kind of finding of a solution to moral dilemma situations as if one would be involved in solving a mathematical problem, but a question of emotional involvement and a warm sense for the needs of one's neighbour (Mk 12:31). Of course, this does not mean that reason becomes superfluous.

> Reasoning frees us from the tyranny of our immediate impulses by allowing us to serve values that are not automatically activated by what's in front of us. And yet, at the same time, reason cannot produce good decisions without some kind of emotional input, however indirect. (Greene 2013, 137)

Such insights about the decisive role of emotions for moral decision-making support also the assumption that good social relationships, like, e.g.,

friendship or attachment within a family, can serve as a kind of training ground for developing adequate moral judgment and play an important role in becoming more sensitive to unjust or harmful behaviour. The emotional attachments we feel for friends or loved ones make us clairvoyant, sensitive, and of course, more vulnerable! Unfairness and fraud among friends hurt much more than among strangers. On the other hand, a society of single individualists who keep polite distance can function very well on a professional level and become cold and inhuman at the same time. Everybody would agree that care provided in a retirement home is not only a question of whether each occupant of the home has received equal portions of the best quality food and is treated according to the best medical standard available. If no personal commitment of the staff and dedication to the care of the old people are present, such "moral correctness" would not be sufficient to ensure a good evaluation from an ethical point of view.

On an anthropological level, the studies of Damasio and his colleagues about the role of feelings and emotions in relation to practical reason support a holistic approach to being human: There is no body-mind-dualism, but a bodily reality, which is an integral part of moral decision-making and moral behaviour. Moral decisions have a physical basis. Here is a linkage to the traditional concept of the soul as "anima forma corporis" by Thomas Aquinas, based on the idea of Aristotelian hylomorphism. The soul as the life principle, which governs all animate things (*anima vegetative*, *anima sensitiva*, *anima intellectiva*), permeates and constitutes the corporeal that is not even conceivable without it. Being human is a natural unity. Soul and body are not two realities or two substances, which subsequently enter into a unity. The reality of the body is the reality of its soul, which requires a physiological and neural substrate to express and represent itself in a given space-time, precisely as a body. One day it might be possible to construct a machine that will solve moral dilemma situations and develop a basic set of adequate reactions to social requirements as politeness, friendliness, and even emotional neediness, as it is shown in the German movie about a humanoid robot, titled *Ich bin Dein Mensch* (Schrader 2021). However, the emotional basis required to perceive moral problems at all presumes a physical reaction to reality: The sudden emergence of compassion towards a little pig in a stable that is far too narrow

and the subsequent build up of hot anger about factory farming is not only a question of simply implanting some "tit for tat"-algorithm into the human brain. The disembodiment of morality within the current discourse about artificial intelligence is an expression of dualism that reduces human being to only some areas of brain (Feichtinger 2017). To put it in the words of the *Memorandum* 2014: "Without the brain everything is nothing, but the brain is not everything!"

2. The Limitations of Neurosciences and the Contributions of Theological Ethics to Interdisciplinary Discourse

In the following part, I point to three aspects that I consider important for interdisciplinary discourse from the perspective of theological ethics. I regard them as requests or necessities for successful dialogue.

2.1 Embedding Neurosciences into a Social Context

Neurosciences have become a leading science. They are fascinating and often offer to us simple answers to difficult and complex questions. At the same time, they provide us with a kind of moral relief: It was not my daughter, who behaved in a strange way. It was her brain! It was not Phineas Gage who reacted unpredictable and irrational, but his brain! I cause accidents, because I cannot park my car correctly. This is because I am a woman who has a brain with a smaller spatial orientation capacity. My colleague tends to react aggressively from time to time because he has a male brain and is, therefore, simply not able to express his emotions adequately. Many popular publications provide us with such "truths", reducing social and moral behaviour to brain functions. Everybody knows that this is much too simple. The plasticity of the brain, modelled by interactions with the social environment, is well-known (Joel & Vikhanski 2019). Yet society loves dualisms, be it the eternal "female-male-dualism", the "body-soul dualism" in former times or the "body-mind-dualism" today. Dualism goes hand in hand with the seductive idea of controlling and manipulating the body to gain control of human mind. The stimulation

of brain areas and the implantation of artificial intelligence might be revolutionary and immensely helpful in treating physical and psychological diseases. However, seen from an ethical point of view, the idea of moral enhancement might be questionable in itself. If moral enhancement could really contribute to bringing forth "morally better persons", what is a sound definition of "morally better persons"? Are "morally better people" just well-adapted to social conventions, "good boys and girls" without any impulse of aggression, peaceful, polite, hard working? Is a world of such people who follow some implanted standard of moral behaviour desirable at all? Admittedly, that world would not include persons such as Donald Trump, but there would neither be Greta Thunberg nor Martin Luther nor Jesus Christ and not even Mother Teresa – just to mention a small selection of famous persons who wrote history by violating established standards of conventional morality. None of them behaved always politely and friendly and, in fact, all of them even showed some kind of aggressive or strange behaviour from time to time. Moral and personal identity are so interwoven with each other that the idea of enhancement or even production of an improved morality could soon become the nightmare of a moral cyborg, equipped with "the latest artificial intelligence-tool" that fosters reciprocity or never-ending patience. Even if there were such a thing like a moral cyborg, we would still have to deal with the question how will he/she react to the surprising challenges of reality, which can never be fully controlled or predicted. Even if she/he had a built-in "turn-the-other-cheek-program", this "love-thy-enemy mechanism" could turn out to be a terrible mistake in a specific situation and increase violent behaviour instead of ending it by setting clear limits.

2.2 Respecting the Limits of Different Perspectives and Methodologies

It is quite evident that theological ethics and neurosciences do not approach reality from the same perspective. It is remarkable that, in order to cross borders, some representatives of neurosciences are not even afraid of what philosophy calls category mistakes. They talk of "dialogues between brains", "feelings of brains", "inventions and compositions" of brains (Werbik & Benetka 2016, 39). This entails the idea of an autonomous

neural being that presents itself as a kind of new human-like moral subject, fascinating and endangering at the same time. In some book titles the brain even becomes "the producer of the soul", a kind of God-like creator (Roth & Strüber 2014). Philosophical and theological ethics have to reject such reductionist super elevation of the brain. Neurosciences describe reality from the observer's point of view. They try to reach conclusions on the basis of imaging procedures and further empirical methods. They study causal relationships, efficiency, and functionality.

In contrast, theological ethics starts from a participant's perspective and claims to be a hermeneutical and prescriptive science, asking about the sense of actions, the good life for everybody, and justice. In the eyes of neurosciences, the story of Phineas Gage is interesting, insofar as they could show which areas in his brain had been affected by the accident and what kind of behaviour was its result. In the eyes of theological ethics, it would be more interesting to know how Gage himself coped with his situation and how his environment reacted to his accident. Maybe the story of Phineas Gage is more the story of his employers, his friends, colleagues, and physicians. At the very least, there is an alternative version of the Phineas-Gage-narrative, recounting how after his employers had dismissed him, he worked as an exhibit (!) in a museum, as a kind of scientific curiosity to the public. Later on, he found a new job and seemed to recover slowly from his accident, while gaining back what his companions named "the old Phineas". Whatever "the truth of this story" might be, searching the world-wide-web still proves a far-spread myth, inspired by sparse facts (Macmillan 2008, 828–831). In any case, this story gives no reason to believe that neuroscience has discovered the full truth about what it means to be human.

2.3 Defending the Concept of Free Will and Moral Responsibility on the Basis of a Relational Anthropology

The *Manifesto of Neurosciences* in 2004, subscribed by leading German neuroscientists such as Gerhard Roth and Wolf Singer, proclaimed, however, the end of the "freedom of will" and cited Benjamin Libet's experiments and those of his successors. 10 years later, the authors of the *Memorandum* conceded that there are

> weaknesses in the field of neuroscience theory, and, on the other hand, in naturalistic presuppositions and concepts that are not well thought out, which make it difficult to build desirable bridges to psychology, philosophy, and cultural studies in the long run.

The ontological reductionism of human persons to their neural system had been all too obvious and the manifest had turned out to be a good example for what happens when neuroscientists naturalize their personal definition of morality or freedom of the will. Each of them has an individual interpretation of the classical philosophical definition of the freedom of the will. According to the common definition (Werbik & Benetka 2016, 62), free decision has to meet three criteria: There must be an option of choices and a real alternative to decide otherwise (the criterion of alternative options). The decision has to depend on the person who decides and not on someone else (the criterion of origination). How the person decides must be subject to her own control, not caused by compulsion (the criterion of control). Besides the necessary fundamental critique, that this classical definition is purely individualistic and neglects relational aspects completely, there has been a long debate in philosophy whether this definition is at all compatible with the findings of neurosciences: The representatives of (1) "incompatibilism" tend to claim that neurosciences have indeed solved the question of free will: There is none. Neuronal activity determines our decision-making and action totally, and the talk of free will is nothing but an illusion. We are neither the authors of our actions nor do we possess free will. The findings of neurosciences question especially the criterion of origination. The representatives of (2) "compatibilism" do not agree. They refer to classical authors like Thomas Hobbes and David Hume and defend the freedom of will in the sense of voluntariness: Even if nature provides human beings with all kinds of desires, freedom consists in autonomy, i.e., the ability to relate to one's desires, to identify with them, or to take a distance from them (Frankfurt 1989, 63–76). I will come back to this influential compatibilist approach later. It entails a rich concept of human agency, leaving space for the assumption that it is not only unconscious mechanic and bodily processes in the brain, but also conscious mental states, that play a role for our decisions and actions.

A third approach, the (3) "libertarian approach", claims that freedom is not only the opposite of determinism but also of pure accident. It is I who is deciding and acting and my decision and act are free if I have good reasons to make a clear choice (Werbik & Benetka 2016, 64).

The last two approaches are interesting for theological ethics, because they leave the mere observer's perspective and choose the participant's perspective. From the perspective of the participant, the freedom of will is an inescapable experience of everyday life (Quitterer 2006, 49). Werbik & Benetka argue that even neuroscientists share this experience and seem to believe in it, otherwise they would not construe neuroprostheses which require the will of patients to move their artificial limbs (Quitterer 2006, 63).

Neuronal processes are a necessary but not sufficient component for decision-making: The capacity to deal and to cope with restrictions, which are part of human contingency and an expression of how to deal with limited freedom, is not only a question of neuronal activity. I remember an impressive story of a patient, told by the psychologist Gerhard Benetka (Vienna): One of his male patients was suffering from Tourette's syndrome. This syndrome is a disorder that involves repetitive movements or unwanted sounds (tics) that are beyond control of patients. They repeatedly blink their eyes, shrug their shoulders, or blurt out unusual sounds or even offensive words. Although there is no cure for Tourette's syndrome, treatments are available. However, in this special case there was a problem: The patient was a highly gifted drummer in a band. It turned out that the medicine he had taken to control his tics made him tired and severely affected his ability to play drums in a club during a weekend. As he was depressed about these side effects, he worked out a plan with his physician: He took his medicine during the week just in order to be able to work as a simple employee. During weekend, he stopped taking it and became the creative drummer, accepting his tics and integrating them into his music. One can hardly consider this acceptance and coping with limitations to be a proof of determinism. It is probably just a good example that participant and observer's perspective can complement each other. In fact, it is possible to deny the freedom of will in theory, but in everyday practice, one cannot but to appeal to the free consideration of someone else to get his/her approval. Even if one becomes a radical determinist after reading Gerhard

Roth and Wolf Singer, he/she would prove – simply by conversion – that there is no determinism, although this argument is, of course, not the most convincing. There might still be a possibility that the brain of the convert has been caught up in a big inescapable illusion. Maybe the whole idea of moral enhancement is such an illusion.

As to representatives of theological ethics, they tend to favour a compatibilist or even libertarian approach. For instance, Stephan Ernst, in order to defend the concept of free will, has rejected the idea that freedom is simply a matter of accident or groundless arbitrariness and has insisted on the existence of reasons that enable free decisions. Freedom of will becomes possible because we have good reasons to make decisions (Ernst 2009, 247). Good reasons do not cancel free decisions but make them possible. Ernst refers to Thomas Aquinas and his distinction between "libertas exercitii" and "libertas specificationis" (S.th. I–II, q. 10, a. 2). The first form of freedom is the fundamental *exercitium* of will – the fact *that* I want (Ernst 2009, 248f). The second one is the specific definition of the action of will, *what* I want. According to Ernst, Thomas goes even beyond these two concepts of freedom and talks about a fundamental freedom, in German a *Grundfreiheit*: The will stands in freedom *vis-à-vis* each of the goods available to choice. For Thomas, the last final justification of this will lies in God. God created human being as oriented towards the Good. Human being has the capacity to know that the good she/he wants to choose is relative and limited. Only the absolute good could move the will with necessity. In reference to a finite good, the choice is not compulsory; the agent can affirm or deny it, because there is an underlying fundamental freedom. Thus, moral judgment is *linked* to neuronal activity, but cannot be *reduced* to it The fundamental freedom Thomas is talking about is compatible with Damasio's insight about the role of emotions. Sometimes they might prevent us from simply following our wishes by a bodily reaction and remind us of the fundamental freedom we have to decide differently. Karl Rahner deepened Thomas's concept of fundamental freedom. Beyond all the different single acts we decide to do, there is the fundamental option in the heart of a person, implying a quest for the Good. It is not the sum of acts but the transcendental moment underlying all these single acts we choose to do (Rahner 1976, 102–105).

Here, at this level, we make the fundamental decision about what kind of person we want to be, although we are often not even explicitly aware of it. This fundamental option manifests itself in many individual acts and attitudes. It is linked to the brain and-at the same time-goes much beyond. Here are the limits of neuroscience. From the perspective of the observer, the fundamental option is not an empirical object of description and is often not even accessible to the participant's consciousness. Here is a story neurosciences can*not* tell us; neither can this story be improved by stimulations of brain areas determined by others. Here is the secret of a person concerning her/his self-determination and her/his fundamental option for the Good. The theological concept of freedom to determine oneself by this fundamental option points to the importance of an anthropology characterized by relationality and responsivity. The awareness of this dimension of human existence forbids theological ethics to think of the moral subject as an isolated brain that needs external stimulation by scientists to behave morally.

3. Conclusions for Moral Responsibility

So far, there has been no empirical evidence that brain processes completely determine social and moral behaviour. Our actions are not merely the result of unconscious mechanistic processes in the brain. Undoubtedly, the latter are "part of the pathway" and may initiate actions (Glannon 2011, 7). However, this is not the whole story of human agency. A rich concept of human agency includes also conscious mental states that play a causal role for forming intentions and executing them in actions. These mental states are shaped by social relations, be it close social relations with family and friends, or by the social context in a wider sense of culture and history. Such concept of human agency confronts us with normative claims and values and demands our answer. It is exactly this responsitivity that characterizes human being from the perspective of theological ethics. Ultimately, for Christian belief it is rooted in the relationship with God. Thus, "more than just the brain is involved in what I choose and do" (Glannon 2011, 7) and what I am responsible for in my actions, my attitudes, and

my fundamental option. Neural processes do not determine free will completely, and the argument of free will being nothing but an illusion is weak (Glannon 2011, 11–13). To sum it up, there is still enough space to hold on to the existence of moral responsibility as a basic idea, as theological ethics has always done. However, there is enough empirical evidence from neurosciences to concede that moral responsibility as a personal capacity is a question of degree rather than a question of "Yes" or "No". Neurosciences can prove that there are forms of limited or reduced capacity of moral responsibility due to accident, cognitive impairment, trauma or simply youth. Obviously, there are efforts in psychiatry and psychotherapy to increase moral and social capacities, but they cannot refrain from social and cultural debates on what morality is and what it should be. The morally perfect cyborg, created and enhanced by neuroscientists, has not been constructed yet, and there are no signs so far that he/she (?) will ever exist. In any case, in the context of a plural society he/she (?) might be a nightmare rather than an achievement.

Bibliography

Crone, E., Das pubertierende Gehirn. Wie Kinder erwachsen werden, München 2016.

Damasio, A., Ich fühle, also bin ich. Die Entschlüsselung des Bewussstseins, Berlin ^{6}2006.

Damasio, A., Descartes' Irrtum. Fühlen, Denken und das menschliche Gehirn. Aus dem Englischen von Hainer Kober, München 1997.

Ernst, S., Grundfragen theologischer Ethik. Eine Einführung, München 2009.

Feichtinger, C., Moral technologies und die Frage der Emotionen. Eine Weiterführung von Peter Kirchschlägers Kritik (2017), https://www.feinschwarz.net/moral-technologies-und-die-frage-der-emotionen-eine-weiterfuehrung-von-peter-kirchschlaegers-kritik/ (06.12.2021).

Frankfurt, H., Freedom of the Will and the Concept of a Person, in: Christman, J. (ed.), The Inner Citadel. Essays on Individual Autonomy, New York/Oxford 1989, 63–76.

Glannon, W., Neuroscience, Free Will, and Moral Responsibility, in: Glannon, W., Brain, Body and Mind. Neuroethics with a Human Face, Oxford 2011, 1–36.

Greene, J. D., Solving the Trolley Problem, in: Sytsma, J., Buckwalter, W. (eds), A Companion to Experimental Philosophy, New Jersey [1]2016, 175–178, https://projects.iq.harvard.edu/files/mcl/files/greene-solvingtrolleyproblem-16.pdf.

Greene, J., Moral Tribes. Emotion, Reason, and the Gap between Us and Them, New York 2013.

Joel, D.,Vikhanski, L, Gender Mosaic. Beyond the Myth of the Male and Female Brain, New York 2019.

Macmillan, M., Phineas Gage. Unravelling the Myth (2008), https://thepsychologist.bps.org.uk/volume-21/edition-9/phineas-gage-unravelling-myth (16.12.2021), 828–831.

Monyer, H. *et al.*, Manifest about Present and Future of Brain Research (2004) = Original source in German: Monyer, H. *et al.*, Elf führende Neurowissenschaftler über Gegenwart und Zukunft der Hirnforschung, "Gehirn & Geist", 2004, 6, 4, 30–37, http://www.hoye.de/hirn/lieferung4.pdf.

Nussbaum, M. C., Politische Emotionen. Warum Liebe für Gerechtigkeit wichtig ist, Frankfurt a.M. [1]2016.

Pascual, L. *et al.*, How Does Morality Work in the Brain? A Functional and Structural Perspective of Moral Behavior, "Front Integr Neurosci", 2013, 7, 65, 1–8, https://www.frontiersin.org/articles/10.3389/fnint.2013.00065/full.

Rosenberger, M., Frei zu leben. Allgemeine Moraltheologie, Münster 2018.

Rahner, K., Grundkurs des Glaubens. Einführung in den Begriff des Christentums, Freiburg i.Br. 1976.

Roth, G., Strüber, N., Wie das Gehirn die Seele macht, Stuttgart [7]2017.

Tretter, F. *et al.*, Memorandum of Reflexive Neurosciences (2014) = Original source in German: Tretter, F. *et. al.*, Memorandum 'Reflexive Neurowissenschaft' (2014), http://www.exp.unibe.ch/research/papers/Memorandum%20Reflexive%20Neurowissenschaft.pdf (02.21.2021), 1–13.

Werbik, H., Benetka, G., Kritik der Neuropsychologie. Eine Streitschrift, Gießen 2016.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 65–79
DOI: 10.24989/BCE.neuro.4

Moral Agents, Brains, and Moral Enhancement

Moralische Akteure, Gehirne und moralische Verbesserung

Elisabeth Hildt, Chicago

Abstract (Deutsch)
In diesem Beitrag geht es um moralisches Verhalten und moralische Verbesserung durch biomedizinische Interventionen. Er beginnt mit einigen Überlegungen zu Moral, moralischen Akteuren und moralischem Verhalten sowie dem wachsenden Bewusstsein für die Rolle des Gehirns für menschliches Verhalten. Anschließend werden das Konzept der moralischen Verbesserung (moral enhancement) vorgestellt und ethische Fragen der moralischen Verbesserung durch biomedizinische Interventionen diskutiert. Darauf aufbauend argumentiert der Text, dass der Ansatz der moralischen Verbesseerung auf einer gehirnzentrierten Perspektive basiert und von dieser angetrieben wird, die einen breiteren Kontext moralischer Handlungsfähigkeit ignoriert.

Abstract (English)
This contribution is about moral behaviour and moral enhancement using biomedical interventions. It begins with some reflections on morality, moral agents, and moral behaviour, and the growing awareness of the role of the brain for human behaviour. Then, the concept of moral enhancement will be introduced, and ethical issues of moral enhancement using biomedical interventions will be discussed. Building on this, the text goes on to argue that the approach of moral enhancement is based on and driven by a brain-centric perspective that ignores the broader context of moral agency.

Keywords (Deutsch)
Ethik; moralischer Akteur; Neurowissenschaft; Gehirn; Neuro-Essentialismus; Neuroethik; moral enhancement; moralische Verbesserung;

Keywords (English)
ethics; moral agent; neuroscience; brain; neuro-essentialism; neuroethics; enhancement; moral enhancement;

1. Humans as Moral Agents

As humans, we see and understand ourselves as persons, as moral agents who are capable of following moral norms. Moral norms refer to morality, a term that can be understood descriptively and normatively. Descriptively, "morality" describes a set of behavioural rules put forward by a society or group or accepted by an individual for their own behaviour. Used normatively, the term refers to a set of rules that would be accepted by every person who meets certain standards, including rationality (Gert & Gert 2020). Traditionally, a person who meets these standards is considered a moral agent.

For the purpose of this text, I stick to the following, very general and comprehensible definition of "moral agent":

> A moral agent is a person who has the ability to discern right from wrong and to be held accountable for his or her own actions. Moral agents have a moral responsibility not to cause unjustified harm. Traditionally, moral agency is assigned only to those who can be held responsible for their actions. (https://ethicsunwrapped.utexas.edu/glossary/moral-agent)

A broad spectrum of ethical theories has been put forward to reflect on morality and build a coherent and consistent network of moral norms. Whereas deontological and consequentialist positions focus on the question of what the right course of action is, virtue ethics positions focus on the moral character of the moral agent. Central to deontological ethics approaches is the nature of an action and whether it adheres to rules or is in

line with duties. Examples are the duty to keep promises, the prohibition against killing, respect for human dignity, or justice. Consequentialism focuses on the consequences of an action: A morally right action is an action that results in good consequences. In this, the criterion for rightness or wrongness is the non-moral value created by an action, e.g., the amount of happiness, financial gain, or health promoted. The proper course of action is the one the consequences of which maximize the non-moral value considered relevant. In contrast, for virtue ethics approaches, moral virtues are central, such as prudence, temperance, or courage. Virtues are about a person's character traits and personality, and relate to their disposition to do what is morally right. A virtuous character is built through continuous exercise throughout life.

We learn from our familial, social, and cultural contexts about what are considered acceptable forms of behaviour. While there certainly are central values, principles, and norms that are valid in almost any societal context, there are also differences in morality, values, and value hierarchies in different societies and groups. Parents who educate their children in various implicit and explicit ways about moral norms have a central role. In addition, there is a broad spectrum of more formal options to positively influence social and moral behaviour, including public encouragement of moral reflection, mentoring, policies, sanctions, and legal regulations.

In general, being a moral agent and trying to do the right thing implies knowing about moral rules and morality as well as the general conditions of the respective situation. It also requires certain capabilities such as rationality, ability to focus, self-control, decision-making capabilities, and motivation. Undoubtedly, the human brain is central for any form of moral behaviour.

2. The Role of the Brain

Facilitated by the so-called "decade of the brain" (1990–1999), neuroscientific knowledge has considerably increased during the past few decades, leading to numerous medical treatments, new forms of diagnosis, and medicine-related technologies. Also, in non-medical contexts, the

awareness of the role the brain plays in health, individual well-being, and human behaviour has increased. Brain-inspired art and new fields like neuro-marketing, neuro-economics, and neuro-education are all indicative of an increased societal role of the brain, a development that has been called 'neuroscientific turn' or 'neuro-turn' (Littlefield & Johnson 2012; Leefmann & Hildt 2017). To give a quantitative example, a study found that the increased attention to the role of the brain is reflected by a 5.5 times increase[1] in the annual frequency of the word "neuroscience" or "neuroscientists" in English language newspapers between 1985 and 2009 (Reiner 2011).

Undoubtedly, the brain has gained relevance for how we see ourselves. A perspective called "neuro-essentialism" may be considered indicative of this. Peter B. Reiner (2011) has characterized neuro-essentialism as the position that we *are* our brains. He writes that neuro-essentialism implies that "when we conceive of ourselves, when we think of who we are as beings interacting in the world, the *we* that we think of primarily resides in our brains." (Reiner 2011, 161).

Similarly, following their analysis of print coverage of neuroimaging, Racine *et al.* (2010) have described interpretations that see the brain as the self-defining essence of a person, as revealing genuine features of the individual, as neuro-essentialism. They have characterized a related tendency as neuro-realism, which they describe as the tendency to interpret brain activation patterns as the ultimate proof that a phenomenon is real and objective (Racine *et al.* 2010).

In line with the growing role of neuroscience in society, research into the ethical and social implications of neuroscience and the role of the brain for social and moral behaviour gained pace, which prompted the development of the field called "neuroethics".

Adina Roskies (2002), when characterizing the field of neuroethics, distinguished between what she called "ethics of neuroscience" and "neuroscience of ethics". Whereas the first subfield of neuroethics examines the ethical and social implications of neuroscience and its medical applications, the subfield of "neuroscience of ethics" investigates central concepts

[1] Normalized to the occurrence of the word "biology".

of philosophical reflection such as free will, self-control, moral cognition, and moral behaviour from the perspective of brain function.

Research in the field of "neuroscience of ethics" deals with a very fundamental tension: While humans as moral agents consider themselves responsible for their actions, the human brain, which certainly is of central relevance for moral behaviour, is a biological organ, subject to the laws of nature. The functioning of the brain depends on neurons, neural activity, neurotransmitters, and receptors, to name just a few of the relevant factors.

For example, in her book "Braintrust. What Neuroscience Tells Us about Morality", Patricia Smith Churchland stresses the role of oxytocin and arginine vasopressin for attachment, social behaviour, and moral behaviour (Churchland 2011). She argues that moral behaviour and morality have their origins in the neurobiology of attachment and bonding. On page 191, she writes (Churchland 2011, 191): "Morality seems to me to be a natural phenomenon – constrained by the forces of natural selection, rooted in neurobiology, shaped by the local ecology, and modified by cultural developments."

Overall, the biological basis of brain functioning seems to be at odds with our perception of being free moral agents.

3. Improving Moral Behaviour – Moral Enhancement

Recently, the idea of aiming to improve moral behaviour with biomedical interventions has gained attention. The currently speculative approach has been called moral enhancement or moral bioenhancement.

The term enhancement originally was introduced in the 1990s to denote procedures as outside the realm of medicine, and thus implies a distinction between medically justified treatments on the one hand and interventions in otherwise healthy individuals, i.e., enhancements, on the other hand (Frankford 1998; Juengst 1998).

The term has also been used in a broader sense to characterize interventions that seek to augment a person's physical or mental capabilities (Lebedev *et al*. 2018). For example, David DeGrazia defines enhancement

(DeGrazia 2014, 361) as "any deliberate intervention that aims to improve an existing capacity, select for a desired capacity, or create a new capacity in a human being." This type of definition also includes medical therapies whose goal is to improve capabilities.

Enhancements include a broad spectrum of interventions that aim at improving human characteristics or capabilities, such as cosmetic surgery, doping in sports, or the use of drugs to improve mental performance. Enhancement interventions aim at improvements, independently of whether these improvements can actually be brought about or not.

With regard to enhancements of mental characteristics or capabilities, a distinction has been made between cognitive enhancement, mood enhancement, and moral enhancement. Cognitive enhancements are interventions that aim at improving a person's cognitive capabilities, such as attention, memory, or concentration. For example, psychostimulants such as Ritalin or amphetamines have been used in an attempt to improve academic performance (Farah *et al.* 2014). Mood enhancement characterizes the strategy of aiming to improve a person's mood, making a person feel better or happier. For example, selective serotonin reuptake inhibitors (SSRIs) have been discussed as mood enhancers (Schermer 2015). Moral enhancement refers to the currently hypothetical idea of improving a person's moral capacities, moral decision-making and moral behaviour.

While scientific evidence for real improvements in cognitive enhancement and mood enhancement approaches is minimal to non-existent, moral enhancement can clearly be characterized as purely speculative at this point.

David DeGrazia (2014) distinguishes between three kinds of moral improvement: a) motivational improvement, i.e., increase the motivation to do what is right; b) increase the understanding of what is right; and c) behavioural improvement, i.e., increase conformity to appropriate moral norms, i.e., do more often what is right.

Even though this distinction is somewhat artificial, it may raise awareness of the various aspects involved in moral decision-making and behaviour.

Based on characterizing the moral status quo as deeply troubling, several authors have fervently argued towards the idea of moral enhancement

using biomedical interventions, i.e., moral bioenhancement. They argue that given widespread injustices, immoral behaviour, poverty, discrimination against minorities, illegal behaviour of all sorts, wars, and terrorist threats, moral enhancement could be a promising and powerful strategy to prevent malicious action and improve or even save mankind (Douglas 2008; Harris 2011; DeGrazia 2014; Gibson 2021).

While these overly negative scenarios of the moral status of humankind seem questionable, the idea of attempting to improve moral behaviour may still be worth considering. As will be discussed below, however, the suggestion of moral bioenhancement is poorly conceived and comes with a number of substantial ethical issues. These will be delineated in the following section.

4. Unresolved Ethical Issues around Moral Enhancement

4.1 What Would Count as Moral Enhancement?

Moral enhancements have been characterized as interventions that aim to improve moral capacities. Examples could be to increase sympathy, increase empathy, increase rationality, reduce aggression, or increase predisposition to altruism or fairness. For example, moral bioenhancements have been suggested to enhance the quality of love in relationships, strengthen marriage bonds, or reduce aggression in perpetrators (Savulescu & Sandberg 2008; DeGrazia 2014). A number of pharmacological substances have been discussed as potential moral enhancers, including oxytocin, dopamine, vasopressin, selective serotonin reuptake inhibitors (SSRIs), methylphenidate, methylenedioxymethamphetamine (MDMA) (Macpherson *et al.* 2019).

It is difficult to clearly describe or specify what modifications in individual characteristics could benefit a person's capability of moral decision-making or moral agency. Besides aspects that relate to rational decision-making, traits like courage, patience, or prudence, i.e., characteristics or dispositions that have been characterized as virtues, come to mind. From this perspective, moral enhancements could consist of interventions

that make people more caring, more patient, or more willing to bond and engage in relationships. These aspects have been stressed by various authors. However, it is obvious that being too patient, too caring, or too relaxed, may come with a number of negative implications. It seems that the optimum may be somewhere in the middle.

In addition, what type of modifications would be considered moral improvements is clearly dependent on many factors. Contexts differ. What may be considered useful in one context might be harmful in another. Value systems differ, in particular in pluralistic societies in which individuals have differing views and value hierarchies. Persons and their individual characteristics differ. A modification in a certain characteristic may turn out to be beneficial in one person, but not in another. And expectations towards a person's behaviour differ. In view of all these factors, it is certainly not clear what sort of modification would count as moral enhancement. Too many factors that differ between individual persons, their life situations, and the contexts in which they interact with others play a role here.

Some authors have attempted to specify in more detail what could count as moral enhancement. In his article "Moral enhancement", Thomas Douglas (2008) emphasizes the role of motives like legal reasoning, love, or sympathy, i.e., factors that may motivate a person to act in a certain way. Being aware that it is difficult to see how an increase in legal reasoning, love, or empathy per se may count as moral enhancement, Douglas (2008) focuses on counter-moral emotions. He characterizes them as emotions "which may interfere with all of the putative good motives (moral emotions, reasoning process, and combinations thereof) and/or which are themselves uncontroversially *bad* motives." (Douglas 2008, 231). Examples he gives are "a strong aversion to certain racial groups", and "the impulse towards violent aggression" (Douglas 2008, 231). He then goes on to think about hypothetic biomedical moral enhancements to alter a person's psychology, the only effects of which supposedly are to improve the person's motives.

This suggestion to eradicate counter-moral emotions with biomedical interventions reveals a high level of brain-centric, if not neuro-essentialist, thinking that assumes that what is central about moral behaviour is

the human brain. Therefore, if one wants to improve behaviour, a brain intervention is called for. Instead, as John Harris rightly points out, what Douglas calls counter-moral emotions may very well be based on false beliefs and prejudice, and a combination of rationality and education may help to get rid of them (Harris 2011).

In addition, Douglas' position seems to involve stereotyped thinking that conceives humans and their brains as being built of something like jigsaw puzzle pieces that can miraculously be exchanged without modifying anything else. It is questionable to assume that certain specific components of "moral behaviour" can be selectively improved without modifying any other of a person's mental capabilities or characteristics. At the very least, potential side effects and risks have to be considered, as interventions are never absolutely safe and effective.

4.2 Moral Enhancement for Whom? – On the Pathologization of Moral Behaviour

Would moral enhancements be something people would want to have for themselves? Or would it be something others would want certain people to undergo, for this or that reason?

Undoubtedly, for each and every person, there is room to improve their moral behaviour, become a better person, be more considerate and supportive of others, and facilitate justice. There may be contexts where becoming "morally better" may be beneficial for the persons themselves; for example, when they repeatedly run into confrontations or arguments with others because of "being maladapted" or "behaving unusually". Sometimes it could be the persons themselves who may want to change their own behaviour. Sometimes, however, it may as well be their fellow human beings that demand change. In the debate around moral enhancement, there is a clear tendency to assume that at least in part, persons should undergo moral enhancement for the benefit of others, i.e., undergo some form of neurointervention so that they are better in line with others' expectations.

Unlike cognitive enhancement, which is perceived as being to the advantage of the individual undergoing enhancement, moral enhancement is

conceived to be an intervention that is motivated at least to a considerable degree by the (presumed) benefit of society.

Several authors have stressed the idea of using moral enhancements in individuals who, for some reason or another, are considered to display inadequate social or moral behaviour (Douglas 2008; DeGrazia 2014; Gibson 2021). For example, to improve the moral behaviour of those displaying impulses towards aggression, or of criminals or terrorists. And by doing so, to improve society.

One question that comes up in this context is: Who would decide for whom moral enhancement would be considered adequate, suggested, or required? And who is it to determine what is the "right" morality?

David DeGrazia (2014) considers public policies to support moral bioenhancement adequate to help reduce what he calls "moral defects". Among the "defects" mentioned are antisocial personality disorder, sadism, intrinsic delight in cheating others, defective empathy, significant prejudice against others, inability to focus on unpleasant realities, weak will, or susceptibility to temptation. While not explicitly stated, this list of "moral defects" seems to imply a new category of disease-like problematic behaviours, i.e., the "moral defects", that require moral enhancement.

Others have used a medicine-centred approach.

Sarah Carter (2017) argues that moral enhancement could, in certain cases, be considered medically indicated. Focusing on psychopathy and a fictionalized mental disorder called "Moral Deficiency Disorder" that is characterized by a deficit of empathy, she discusses how lack of empathy could be considered a deficiency, something pathological that could make moral enhancement medically indicated and thus therapeutic. This fictionalized perspective strengthens the idea that "medicalisation of morality" or "pathologization of morality" could be an adequate approach.

Richard Gibson (2021) refers to the idea of herd immunity of morality. He argues that, provided a large enough percentage of the population voluntarily used moral bioenhancement, the behavioural equivalent of the herd immunity threshold could be reached so that "ultimate harm", i.e., humanity's self-inflicted destruction, can be avoided. According to Gibson, moral bioenhancement could "interrupt the transmission chain of 'ultra-high-intensity' negative behaviours" (Gibson 2021, 51), by "disrupting

the chain of behavioural infection" (Gibson 2021, 51). Once a minimum threshold is reached, he argues, a significant reduction of the incidence and spread of immoral behaviours could be achieved. By alluding to a vaccination analogy, he seems to imply that morally good behaviour can be injected or applied in some other efficient way, which makes moral behaviour a matter of epidemiology. A perspective on moral behaviour like this seems overly biologistic, if not deterministic; it clearly disregards the complexity of moral behaviour. Also, it seems to assume tacitly that moral bioenhancement is dealing with some sort of disease that is to be eradicated, as revealed by phrases like "removal of the susceptible host from the transmission triad".

4.3 Autonomy and Cognitive Liberty

The "right to cognitive liberty" has been a central pillar in interdisciplinary discussions on enhancement. Wrye Sententia (Sententia 2004, 223) characterized cognitive liberty as "every person's fundamental right to think independently, to use the full spectrum of his or her mind, and to have autonomy over his or her brain chemistry." The right to cognitive liberty stresses every person's right to freely and autonomously decide whether or not to undergo any form of neurointervention or neuroenhancement. However, it is exactly the right to cognitive liberty that was questioned by contributions that suggest using moral bioenhancement for the benefit of society. Infringements on the right to cognitive liberty and the related right to mental integrity would be most obvious if moral enhancements were to be used in prisoners or criminals. If criminals were forced to undergo moral bioenhancements or prisoners asked to give their consent to a moral bioenhancement intervention that implies shortening their time in prison, there clearly would be undue influences that limit voluntariness.

That's why Elizabeth Shaw (2018) invokes the rights to mental and bodily integrity to argue against a mandatory use of neurointerventions in criminal sentencing. Also, David Birks and Alena Buyx (2018) refer to an interest in mental integrity when arguing against administering neurointerventions to offenders as an alternative to incarceration.

Despite this, in their article "The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity", Ingmar Persson and Julian Savulescu (2008) explicitly argue in favour of compulsory moral bioenhancement. They consider moral enhancement an imperative in view of potential threats of cognitive enhancement misuse and the need to prevent dangerous, immoral people from misusing scientific progress towards their own ends.

They write (Persson & Savulescu 2008, 174):

> If safe moral enhancements are ever developed, there are strong reasons to believe that their use should be obligatory, like education or fluoride in the water, since those who should take them are least likely to be inclined to use them. That is, safe, effective moral enhancement would be compulsory.

Another autonomy-related question discussed in this context is how far moral bioenhancement would limit freedom, in that morally enhanced persons simply could not do otherwise than doing the right thing or having the right motives (Douglas 2008; Harris 2011; Diéguez & Véliz 2019). It was argued that freedom, even the freedom to have bad motives, is valuable (Douglas 2008). Whereas freedom certainly is a central component of sound morality and of being a moral person (Harris 2011), it seems doubtful whether moral biointerventions would ever be able to fully determine a person's behaviour. David DeGrazia (DeGrazia 2014) argued that moral enhancement would not threaten freedom, as the person still performs intentional actions that are under voluntary control. In his view, maximal freedom is not necessary; the benefit is to have improved moral behaviour.

5. Conclusion

In the discussion on moral enhancement, moral behaviour is not conceived as a matter of values, beliefs, convictions, rationality, respect for others, discerning right from wrong, or responsibility. Neither is it con-

ceived as a matter of education, enculturation, or socialization. Instead, it is conceived as a matter of brain functioning. If a person's moral behaviour is considered problematic, it must be because of their brain. It is not the knowledge of the nature of an action or the consequences of an action, or the cultivation of moral virtues that the authors think of when they reflect on moral behaviour and how to improve moral behaviour, but brain chemistry.

While it is certainly correct that human moral behaviour is critically dependent on the human brain, the debate on moral bioenhancement relies heavily on a brain-centric, if not neuro-essentialist view. Improving moral behaviour becomes a matter of currently hypothetical neurointerventions called moral bioenhancements that modify brain functioning and thus selectively influence certain motives or character traits. All of this ignores the complexity of moral behaviour and morality and the manifold influences at the individual, social and cultural level.

Bibliography

Birks, D., Buyx, A., Punishing Intentions and Neurointerventions, "AJOB Neuroscience", 2018, 9, 3, 133–143, DOI: 10.1080/21507740.2018.1496162.

Carter, S., Could Moral Enhancement Interventions be Medically Indicated?, "Health Care Anal", 2017, 25, 338–353.

Churchland, P. S., Braintrust. What Neuroscience Tells Us about Morality, Princeton, NJ 2011.

DeGrazia, D., Moral Enhancement, Freedom, and What We (Should) Value in Moral Behaviour, "J Med Ethics", 2014, 40, 361–368.

Diéguez, A., Véliz, C., Would Moral Enhancement Limit Freedom?, "Topoi", 2019, 38, 29–36.

Douglas, T., Moral Enhancement, "J Appl Philos.", 2008, 25, 3, 228–245.

Douglas, T., Forsberg, L., Three Rationales for a Legal Right to Mental Integrity, in: Ligthart, S. *et al.* (eds), Neurolaw, London, 2021, 179–201, https://doi.org/10.1007/978-3-030-69277-3_8.

Farah, M. J. *et al.*, Cognitive Enhancement. Wiley Interdisciplinary Reviews, "Cognitive Science", 2014, 5, 1, 95–103.

Frankford, D. M., The Treatment/Enhancement Distinction as an Armament in the Policy Wars, in: Parens, E. (ed.). Enhancing Human Traits. Ethical and Social Implications, Washington, D.C. 1998, 70–94.

Gert, B., Gert, J., The Definition of Morality, in: Zalta, Edward N., The Stanford Encyclopedia of Philosophy (Fall 2020 Edition), https://plato.stanford.edu/archives/fall2020/entries/morality-definition/.

Gibson, R. B., The Epidemiology of Moral Bioenhancement, "Medicine, Health Care and Philosophy", 2021, 24, 45–54.

Harris, J., Moral Enhancement and Freedom, "Bioethics", 2011, 25, 2, 102–111.

Ienca, M., Andorno, R., Towards New Human Rights in the Age of Neuroscience and Neurotechnology, "Life Sci Soc Policy", 2017, 13, 5, https://doi.org/10.1186/s40504-017-0050-1.

Juengst, E. T., What Does Enhancement Mean?, in: Parens, E. (ed.), Enhancing Human Traits. Ethical and Social Implications, Washington, D.C. 1998, 29–47.

Lebedev, M. A. *et al.*, Editorial: Augmentation of Brain Function. Facts, Fiction and Controversy, "Front. Syst. Neurosci.", 2018, 12:45, DOI: 10.3389/fnsys.2018.00045.

Leefmann, J., Hildt, E., Neuroethics and the Neuroscientific Turn, in: Rommelfanger, K., Johnson, S. (eds), The Routledge Handbook of Neuroethics, Abingdon 2017, 14–32.

Littlefield, M. M., Johnson, J. M. (eds), The Neuroscientific Turn. Transdisciplinarity in the Age of the Brain, Ann Arbor, MI 2012.

Macpherson, I. *et al.*, Moral Enhancement, at the Peak of Pharmacology and at the Limits of Ethics, "Bioethics", 2019, 33, 992–1001.

Persson, I., Savulescu, J., The Perils of Cognitive Enhancement and the Urgent Imperative to Enhance the Moral Character of Humanity, "Journal of Applied Philosophy", 2088, 25, 3, 162–177.

Racine, E. *et al.*, Contemporary Neuroscience in the Media, "Soc Sci Med.", 2010, 71, 4, 725–733.

Reiner, P. B., The Rise of Neuroessentialism, in: Illes, J., Sahakian, B.J. (eds), The Oxford Handbook of Neuroethics, New York 2011, 161–176.

Roskies, A., Neuroethics for the New Millenium, "Neuron", 2022, 35, 21–23.

Savulescu, J., Sandberg, A., Neuroenhancement of Love and Marriage. The Chemicals Between Us, "Neuroethics", 2008, 1, 31–44.

Schermer, M., Ethics of Pharmacological Mood Enhancement, in: Clausen, J., Levy, N. (eds), Handbook of Neuroethics, Dordrecht 2015, 1177–1190.

Shaw, E., Against the Mandatory Use of Neurointerventions in Criminal Sentencing, in: Birks, D., Douglas, T. (eds), Treatment for Crime. Philosophical Essays on Neurointerventions in Criminal Justice, Oxford 2018, 321–337.

Sententia, W., Neuroethical Considerations. Cognitive Liberty and Converging Technologies for Improving Human Cognition, "Ann N Y Acad Sci", 2004, 1013, 221–228.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 81–94
DOI: 10.24989/BCE.neuro.5

The Craft of Freedom

The Relevance of Peter Bieri's Concept of "Appropriated Freedom"

Das Handwerk der Freiheit

Die Relevanz von Peter Bieris Konzept der „angeeigneten Freiheit"

Walter Schaupp, Graz

Abstract (Deutsch)
Der Artikel versucht, die Theorie der „angeeigneten Freiheit", wie sie der Schweizer Philosoph Peter Bieri in seinem Buch Das Handwerk der Freiheit. Über die Entdeckung des eigenen Willens ausgearbeitet hat, auf aktuelle Probleme anzuwenden. Bieris Freiheitsbegriff, ein kompatibilistischer Ansatz, stellt das Selbst als den entscheidenden Faktor in den Mittelpunkt, in dem sich Autonomie und Selbstbestimmung primär realisieren. Durch den Prozess einer bewussten und reflexiven Aneignung früherer spontaner und „unfreier" Wünsche und Willensäußerungen bildet sich ein neues, „freies" Selbst. Das Konzept hat seine Grenzen, aber es ist in der Lage zu erklären, was Selbstbestimmung angesichts neuer Formen indirekter gesellschaftlicher Einflüsse und manipulativer Kräfte bedeuten könnte. Darüber hinaus verdeutlicht Bieris Theorie der angeeigneten Freiheit die Rolle der spirituellen Suche und Praxis in Bezug auf die Freiheit.

Abstract (English)
The article tries to apply the theory of an "appropriated freedom" as it is elaborated by the Swiss philosopher Peter Bieri in his book *The Craft of Freedom. About the Discovery of One's Own Will* to actual problems. Bie-

ri's concept of freedom, a compatibilistic approach, focuses on the *self* as the decisive factor, where autonomy and self-determination are primarily realized. Through the process of a conscious and reflexive appropriation of former spontaneous and "unfree" desires and volitions a new, "free" self is formed. The concept has limitations but it is able to explain what self-determination could mean in the face of new forms of indirect social influences and manipulative forces. Moreover, Bieri's theory of an appropriated freedom elucidates the role of spiritual search and practice with regard to freedom.

Keywords (Deutsch)
Theorie der Freiheit; Freiheit; Selbstbestimmung; Peter Bieri; Subjektivierung; Spiritualität;

Keywords (English)
theory of freedom; freedom; self-determination; Peter Bieri; subjectification; spirituality;

Peter Bieri is a Swiss philosopher, born in 1944 and still alive, who worked at the Universities of Bielefeld, Heidelberg and Berlin and engaged himself for a long time in analytical philosophy. His book *Das Handwerk der Freiheit. Über die Entdeckung des eigenen Willens* (*The Craft of Freedom. On the Discovery of One's Own Will*) (Bieri 2001) is located within the so-called *mind vs. brain* debate, a highly specialized controversy between neuroscience and philosophy of mind in the decades around the millennium.

1. Between Determinism and Libertarianism

The debate was provoked by neuroscientists claiming that all mental phenomena are completely reducible to neuronal processes. There exists, so the thesis of the "reductionist" or "naturalistic" position in this controversy, no ontologically distinct "mind" or "soul" in humans. The only reality

are neuronal networks of a highly complex nature in our brains which we, as human subjects, *experience* as "mind", "soul" or as a substantial and distinct "I". To conceive ideas like "mind", "soul", or an immaterial "I" as real entities would be a mere illusion, a form of self-deception the conscious "I" has about itself and its true nature.

An important side stage of the debate was a discussion on freedom. Does human freedom exist at all? How should we interpret our undeniably experienced intuition to act and behave freely in the light of the new neurobiological insights? Reductionists or naturalists like Wolfgang Prinz, Wolf Singer, and Gerhard Roth stood up for a strict *deterministic* position, maintaining that behind every act of conscious will and every decision we make a neurobiological event takes place in our brains which triggers the action and is itself triggered by some other biological processes. According to this concept, there is no room for freedom. In contrast to that concept, *libertarians* like Robert Kane, Karl Popper, and Rodrick M. Chisholm believe that there is a distinct, immaterial agent in us, which is not part of the physical or biological structures. This agent can cause actions independently from physical processes, which in this sense are "free will" actions. Such positions presuppose a dualistic world view as they have to assume a reality beyond our physical world in time and space.

A range of theories in philosophy of mind try to overcome this hard dichotomy between naturalism and dualism, between strict determinism and an absolute, unconditioned freedom. The *double-aspect theory*, for example, holds that the mental and the physical are only two aspects ("subjective" and "objective") of one and the same reality which, in itself, is not accessible to our mind.[1]

Compatibilism is another attempt to establish a third way between determinism and unconditioned freedom. Representatives of this position are Michael Pauen, Ansgar Beckermann, and Peter Bieri, whose approach will be discussed in this contribution, although the publications of Beckermann (e.g., Beckermann 2012) have attracted more attention

[1] The account goes back to Baruch Spinoza. In the context of the modern debate, moral theologian Michael Rosenberger defends such an approach in his book *Determinismus und Freiheit* (Rosenberger 2006).

than those of Bieri. Compatibilists hold that a deterministic world view and the assumption of human freedom need not exclude each other but can co-exist. What these approaches have in common is a shift in the notion of freedom which they understand as *authenticity*. It does not primarily imply that a will or an intention within us is not determined by neuronal events but that it is *in congruence with a deeper – again neuronally codified-self* in us. If we can act out what we want, according to our inner self, we are free.

There are two important arguments for such a "weak" understanding of freedom. First, there is the everyday life-experience that we feel free if we are in complete harmony with ourselves and with the world around us. If we do not feel any alienation within us, we feel free. The cell of a contemplative monk is, for example, a very restricted environment and could be seen as a kind of prison. But the monk, who lives in this cell, can consider himself completely free, if he identifies with this setting, if he does not feel any alienation. Secondly, if we detach human acts radically from the "self" which lies behind them, these acts become arbitrary in a way we would not associate with real freedom. Normally, we also assume that free self-determination is in some way connected to the character, the biography, and the inner aspirations of a person.[2]

At any rate, as we will see, for Bieri the central problem of defining freedom also shifts from the question how isolated acts are brought about by the underlying self to questions like how is this self formed, and could it be that the mystery of freedom primarily has to be located here? – After a short presentation of Bieri's concept of freedom and some critical comments on it, the aim of this contribution is to show how his idea of an "appropriated freedom" can shed a new light on therapeutic and pastoral caring processes, on the individual who is subjected to societal conditions, and on the relevance a spiritual life could have for freedom.

[2] The difference between arbitrariness and real freedom is commonly accepted in the philosophy of mind. Also, in a Kantian perspective freedom is not complete in-determination and so arbitrariness but determination of the will by the law of practical reason.

2. Appropriated Freedom

For Bieri freedom is not a quality or a characteristic of human acts which we always have at hand in an uncomplicated way. It is something we have to acquire actively in our life. This is what is meant by the word "craft" in the title of his book. Freedom is realized or acquired through a three-level process of "appropriation" (*Aneignung*) of our desires, volitions, and strivings, which emerge spontaneously within us and are subjectively experienced by us as our "will". In the first instance such a "will" is completely determined by (neuro-)biological mechanisms and chains of reactions, and in this sense is "unfree". By a conscious process of appropriation, it becomes "ours" in an emphatic sense, and now, in living it out, we are free. The three stages are:

1) *Articulation* ("Der freie als der artikulierte Wille"; Bieri 2001, 385): The effort of articulating a "will" or "desire" within us, by spoken language or in written form, equates to the first level of freedom. What has moved us unconsciously, now becomes objectifiable becomes objectifiable – for us and for others, whom we tell about it. Through this we enter into a new, conscious, and reflexive relationship with such *desires* and thereby with our own *self*. We are not dominated by a will in the same way as before, if we now have a clear idea of what exactly is driving us.
2) *Understanding* ("Der freie als der verstandene Wille"; Bieri 2001, 389): On the second level, we begin to understand *why* this will exists within us and has such a dominating influence on us. Reasons and origins can lie in our biography, in certain experiences we have made, and they can be of a genetic, psychological, or sociological nature (genes, character, education, socialisation, life-events). Understanding can also mean gaining insight into what was previously a matter of self-deception. For instance, we may be able to understand now that some tendency within us is not only something good but can also have negative implications that we failed or refused to notice previously. Generally spoken, by understanding why we feel or behave in a certain way, we gain a modified stance towards ourselves and, at the same time, it is our self

which is altered. Again, for Bieri, understanding why we are as we are and feel as we feel equates to a step in the direction of greater authenticity and of freedom. Once we understand something inside of us, which we always have been worried about, we can accept it in a new way.

3) *Approval* ("Der freie als der gebilligte Wille"; Bieri 2001, 397): On the third level, after having articulated and understood a will within us, we can either identify or not identify with it, we can either approve or disapprove of it. In this step, something which has been part of our self before now becomes part of a new identity, part of a reflexive self, which has taken a conscious and reflexive stance on it. On this level it makes a big difference if something determines us against our explicit will or if we act in free accordance with it.

Every such appropriation of a "will" within us leads to experiencing a new self. We can open up ourselves consciously to such processes or we can avoid them. If we avoid them, we remain unfree – at least in this respect. This process has to be undergone again and again and requires energy and concentration. This is what "craft" in the title of the book refers to. Self-determination and autonomy are realized primarily by such processes. However, there is a possibility that we become unfree again, because it could happen that we live a kind of life which we could not approve any longer if we were to re-enter the process of appropriation. But there is also a possibility to grow in authenticity and thus experience freedom in a lasting way.

3. Limits of the Concept

Before outlining possible positive implications and practical consequences of Bieri's theory of freedom some of its limits should be discussed. First of all, one wonders if the mere *articulation* and *understanding* of something (level 1 and 2) can be seen as a realization of *freedom*. Normally we link articulation and understanding to our ability of recognition and freedom to the spheres of will and action. Cognition can be true or false, actions

can be right or wrong. Moreover, not all decisions in our life that we consider autonomous are preceded by an explicit process of articulation and understanding like Bieri seems to require. More radically, if freedom really presupposes adequate understanding, we will never be free, because we can never be sure that we have already reached such adequate understanding of ourselves and of the world.

Obviously, the concept of Bieri is too restricted and reductive in this respect. It is questionable if what he describes is the only form of freedom and self-determination we exercise in our life. If one follows Beckermann, who has been mentioned above, it could very well be that the degree of freedom corresponds to the amount of time we invest in considering and deliberating certain actions.

But Bieri is certainly right in his observation that *understanding* myself in a substantial new way *alters* the "self" that I am, and that such a *transformation of identity* will have consequences for my *attitudes*, *decisions*, and *actions*. It is obvious that there is a thread leading from new *self-explication* to new *self-understanding* and from there to new possibilities of *action*. It is hard to believe that such a process of transformation, if it is realized on a conscious level, has nothing to do with freedom. On the contrary, people can readily close themselves off to new insights and truths can readily be neglected if they question the own identity. Bieri's concept highlights the fact that we exercise freedom in our life on this level in an essential, but easily ignored way. The important message is that it is not sufficient to look at the mental events of will within us and follow physical processes, like it was the case with Benjamin Libet's famous experiments (Libet 1985). Bieri's contribution becomes relevant, insofar as isolated events of will and action are obviously determined and restricted in their range by underlying factors, which have become a part of our self and which we are not conscious of.

There is another important objection to Bieri's theory or freedom. It is related to the third level of appropriated freedom, i.e., to the approval/non-approval of an articulated and understood will within us. Within a strictly deterministic framework, which is accepted by compatibilists, there is the problem of how "free" this mental act of approval/non-approval really is. Is it not necessarily again caused by neurobiological events? This

problem has been discussed extensively with regard to Harry Frankfurt's theory of first and second order desires which obviously stands behind the third-level idea of Bieri (Frankfurt 1971). According to Frankfurt, freedom is grounded in the ability of human agents to generate a higher order will, a so called "second order desire", by which a person takes a stance on a spontaneous will, something that animals are not able to do. However, this proposal would lead to an infinite regress, as we would constantly need desires of a higher level to solve the problem of determinism.

As already mentioned, the aim of this contribution is not to solve the problem of determinism/indeterminism on the ontological level but to examine the possible contribution of Bieri's theory to a better understanding of freedom in the context of neuronal and social determination. Again, there the true core of Bieri's theory is that actions which flow from a desire or a will, which has been thought over for a long time, surely can be seen as more "autonomous" than an instant and impulsive action triggered off by the same will.

Hence, Bieri's concept has an important message, but it has to be modified. We should distinguish between a *constitutive ability* of freedom, which we always have, and the *realisation* of freedom, which can take place on different levels and which requires some work, as Bieri insists. Also, the Catholic moral tradition assumed that freedom and, along with it, personal moral responsibility, are not always realized in the same way. Following St. Thomas Aquinas, one ought to distinguish between "actus *humanus*", an act which is conscious and deliberate, and "actus *hominis*", all acts physically caused by a human being, including those caused in an unconscious, spontaneous, or enforced way (Thomas 1962, 1.2.q.1 aa.1,3).

The idea that humans realize their ability of freedom in different acts in different levels is transferred by Bieri to the level of the self. Our self, and the actions flowing out of it, can be either totally uncontrolled or can be transformed into an "autonomous" self by the process of appropriation described above. For Bieri this is the central task of freedom in our life. It is an approach which stresses the fact that human freedom is radically embedded in the interplay of desires and aspirations which form our natural self. What we can do is to work with all these determining factors.

In a broader perspective this can be seen as a *translation of existential thinking*, which was the dominating philosophy in the middle of the last century, into the terms of analytic philosophy and neurobiology. The central question of existential thinking is: "Who do I want to be?" It is a call to take on the inescapable responsibility for one's own existence. Within theology, it was Karl Rahner, who defended such a concept of freedom. As the following citation shows, it comes fairly close to what Bieri wants to tell us: "In its original sense, freedom is not the capacity to choose this or that object or mode of behaviour, it rather is the freedom of self-understanding, the possibility of saying yes or no to oneself, the possibility of deciding for or against oneself." (Rahner 1962, 223[3])

4. Relevance of the Concept

4.1 Psychotherapy and Pastoral Counselling

Following the concept of appropriated freedom, it becomes obvious that psychotherapy and consultation are not only "healing" practices, but can also be understood as a work of freedom. They are places where we try to articulate inner feelings, desires and thoughts, where we try to understand them in a new way, and where we reflect on new possibilities for action. Though therapeutic assistance itself should be value-free, it is clear that the aim of such processes is to stimulate conscious and deliberate stances on various facets of the own self.

The same applies to pastoral counselling. Being on the way with people and providing them with a room for communication and possible resonance in difficult, oppressing, and confusing situations of their lives can help people overcome their problems and experience a new kind of freedom. In a world of multiple and ongoing social pressures, it is important that the message of the gospel can grow on a ground of freedom and not of coercion and pressure.

---

[3] Translation by the author.

4.2 The Individual and the Powers of Society

Freedom is not only a philosophical and a neurobiological issue, it is a key value of modern society which has inspired many generations before us. At the beginning, the struggle for freedom and emancipation aimed to overcome external, heteronomous forces and powers, specifically in the form of religious and secular authorities. The aim was to get rid of the oppressive influence of persons, institutions, and traditional norms to be able to think independently and to lead a life according to one's own believes and convictions. In the second phase it became clear that freedom means not only the absence of external oppression but also to be able to dispose of the necessary material and societal resources in order to live a self-determined life.

What we are experiencing currently is a *new demystification of the autonomous subject*. Despite the obvious absence of external powers and the growing material resources we do not really feel free and are in fact confronted with new forms of manipulation. Various sociological analyses describe new forces and powers which have taken the place of previous external powers and attempt to control the individual in a more subtle way. They try to form, even manipulate our wishes and to determine our behaviour. This is the case with marketing and commercials (Steyrer 2018; Kahneman 2011), with the many "influencers" and the producers of fake news in the social media platforms, and with market orientation and competitiveness in general. For authors who follow Michel Foucault's line of thinking, the modern individual is submitted to various forms of "subjectification". Various social "dispositifs" bring about certain types of subjects, which feel, think, and desire in a certain way, not knowing why they do so (Bröckling 2016). This is why Judith Butler insists that the modern individual strictly remains opaque for himself/herself and for others. For her, there is no possibility of completely understanding all the inscriptions which have brought about the way we experience and understand the world (Butler 2001 and 2007).

If one wants to respond to these challenges, it is not enough to refer to freedom of will and freedom of action in a straightforward way. If it is true that our self is co-created by such powers, the struggle for freedom

will assume the form Bieri has outlined. We have to become aware of what drives us, we have to articulate and understand the underlying systems, and finally have to develop a reflexive and conscious stance towards it. This does not imply that we will always have to discard what is in us and what drives us. Freedom only requires us to deal with such influences reflectively and reasonably.

4.3 Spirituality and Pastoral Work

Besides that, there are some interesting interrelations and correspondences between Bieri's concept of appropriated freedom and the spiritual life. The concept enables us to understand the nature of spirituality and of certain spiritual practices, partly specific to the Christian tradition, in a new way.

The first point refers to the growing *importance of meditation* and *practices of awareness* and of retreat within the field of spirituality and religion. The object of meditation can be religious texts like the bible, but there is also a growing interest in athematic meditation and various forms of practices of awareness. Against the background of subjectivation, explained above, all this can be understood as a search for an inner space of awareness where we will become aware of the pressures and norms of society that form our self and where we can begin a confrontation with them on a new, conscious level. Obviously, such efforts are an important exercise of autonomy in our life.

The second point concerns the spiritual *doctrine of "discernment of the spirits"*. Its origins lie in the time of the desert fathers and later it has been systematically unfolded especially by Saint Ignatius of Loyola in the sixteenth century. The human mind-in classical texts the human "soul"-is seen as under the influence of good and evil "spirits", inner forces we experience, and the doctrine wants to define criteria by which we can discern them. We are guided to recognize those which will lead to true life and true freedom and to reject the others. For Ignatius in the end there is only one good spirit, the Holy Spirit, who stands behind all forces that drive us towards life, wholeness, and peace. According to Ignatius, in the long run the bad spirit will evoke anguish, confusion, and desolation in us, whereas

the good spirit is marked by inner peace, consolation, joyfulness, and motivation to act (Ignatius 1966, 104–110).

There is an interesting correspondence between this traditional image of human being under the influence of good and bad "spirits" and the modern concept of the individual as a subject which is affected by social discourses, expectations, and norms which are often internalized in a way that makes it difficult to perceive them as foreign and alienating powers. It seems to be clear that in view of this constellation we urgently need a new art of discernment, which, against the background of Bieri's theory, is nothing other than a work of freedom.

The last point makes reference to the *biblical understanding of freedom.* The challenging thing about Bieri's conception of freedom is that there should be a more or less large degree of freedom, not in the sense of more liberal or coercive political circumstances but of freedom as a feature or quality of human existence itself. For Bieri, it is the inner self that is more or less "free" and that has to be set free by freedom work, by a repetitive effort of the conscious appropriation of one's own will. There are analogies to this in the understanding of freedom one can find in the texts of St. Paul and St. John and that shall only be touched upon briefly.

For Paul, the existence of the converted and baptized Christian is marked by a new freedom as an effect of the spirit living in him. As we read in 2 Corinthians 3:17: "Now the Lord is the Spirit, and where the Spirit of the Lord is, there is freedom". Similarly, Galatians 5:1 states: "It is for freedom that Christ has set us free. Stand firm, then, and do not let yourselves be burdened again by a yoke of slavery." In John 8:31-32 this is confirmed: "If you hold to my teaching, you are really my disciples. Then you will know the truth, and the truth will set you free." Here the disciples ask back: "We are Abraham's descendants and have never been slaves of anyone. How can you say that we shall be set free?" On that, Jesus responds: "So if the Son sets you free, you will be free indeed" (John 8:32-33,36). For both, John and Paul, becoming a disciple of Christ means to undergo a transition from (inner) slavery to sin to an existence in the freedom of the children of God. It is a freedom which has not been there before. If, with John, it is the "truth" (the same reality can also be called "light") which makes free, there is, like with Bieri, a connection between freedom and understanding.

The one who has accepted the *message* of Jesus and bears his *spirit* within himself has reached a *new self-understanding* which in turn brings about a *new freedom*. Admittedly, there is one important difference between Bieri's view of becoming free and the biblical one. For Paul and John, it is clear that the new freedom is primarily a gift of grace and not the fruit of human effort. But if one looks at the subsequent understanding of supernatural grace and human freedom in the Christian tradition, it becomes clear that grace and human effort do not exclude each other. Also, from a Christian perspective the new freedom can be seen as an effect of grace as well as the work of human being.

5. Conclusion

The emerging neurosciences have put human liberty into question. Peter Bieri's compatibilistic answer to the problem of neurobiological determinism shifts the centre of attention from the generation of isolated, external human acts to the formation of the human self. Freedom is the ability to engage actively and in a reflexive way with the different factors which influence and form this self and to construct a new self which is more our "own" than it was before. This account not only enables us to regard all forms of psychotherapeutic work as a work on freedom, it also opens a way how we could understand freedom in the face of all the societal powers which take control of us by invading our self-understanding, thinking, and wishing from *within* ourselves. It also allows us to discover a positive link between spirituality and freedom. Contemplative practices open up a new, inner space of awareness for the self and so can play an important role for the modern individual in his/her search for true autonomy.

Bibliography

Beckermann, A., Gehirn und Freiheit, in: Fink, H., (ed.), Verantwortung als Illusion, Paderborn 2012.

Bieri, P., Das Handwerk der Freiheit. Über die Entdeckung des eigenen Willens, München 2001.

Bröckling, U., Das Unternehmerische Selbst. Soziologie einer Subjektivierungsform, Frankfurt a. M. 2016.

Butler, J., Kritik der ethischen Gewalt, Frankfurt a. M. 2007.

Butler, J., Psyche der Macht. Das Subjekt der Unterwerfung, Frankfurt a. M. 2001.

Frankfurt, H., Freedom of the Will and the Concept of a Person, "The Journal of Philosophy", 1971, 68, 1, 5–20.

Ignatius von Loyola, Geistliche Übungen. Übertragung und Erklärung von Adolf Haas, Freiburg i.Br. 91966, 104–110 (EB 313–336).

Kahneman, D., Schnelles Denken, langsames Denken, München 2011.

Libet, B., Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action, "The Behavioral and Brain Sciences", 1985, 8, 529–539.

Rahner, K., Theologie der Freiheit, in: Rahner K., Schriften zur Theologie. Band VI, Einsiedeln 1965.

Rosenberger, M., Determinismus und Freiheit. Das Subjekt als Teilnehmer, Darmstadt 2006.

Steyrer, J., Die Macht der Manipulation, Salzburg 2018.

Thomas von Aquin, Summa Theologiae, Roma: Ed. Paulinae 1962.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 95–109
DOI: 10.24989/BCE.neuro.6

What Can Neurosciences Tell Us?

Brain Injuries and Their Impact on Behaviour

Was können wir von den Neurowissenschaften lernen?

Hirnverletzungen und ihre Auswirkungen auf das Verhalten

Petr Hluštík, Olomouc

Abstract (Deutsch)

Es ist seit langem bekannt, dass Hirnverletzungen und -schädigungen das menschliche Verhalten beeinflussen können. Der vorliegende Text beleuchtet einige ausgewählte Beispiele für sozial unangemessenes, gewalttätiges oder sogar schwer kriminelles Verhalten bei Menschen, die Hirnschäden erlitten haben, am häufigsten im Frontalhirn. Erstens geht es um das bekannteste Exempel aus dem Lehrbuch, nämlich um den amerikanischen Eisenbahnkonstrukteur Phineas Gage, dessen Gehirn 1848 verletzt wurde. Die daraus resultierende Verhaltens- und Persönlichkeitsveränderung wurden von seinem Arzt beschrieben und anschließend in vielen, wenn nicht den meisten Lehrbüchern der Neurologie, Psychologie und Neurowissenschaft zitiert. Eine weitere, neuere Überlegung betrifft die Feststellung, dass Straftäter, die in den USA zum Tode verurteilt werden, häufig an unerkannten schweren psychiatrischen, neurologischen und kognitiven Störungen leiden. Diese sind für Überlegungen zur Strafmilderung relevant. Schließlich kann Aggressivität bis hin zum Mord ein (seltenes) Merkmal von Infektionskrankheiten des Nervensystems sein, wie z. B. der Borreliose. Diese Beispiele werfen viele Fragen auf: Inwieweit kann moralisches (oder unmoralisches) Verhalten durch eine Hirnfunktionsstörung beeinflusst werden? Kann das dokumentierte Vorhandensein von Hirnschäden

und -funktionsstörungen kriminelles Verhalten teilweise oder ganz entschuldigen?

Abstract (English)
It has long been known that brain injuries and lesions may influence human behaviour. The present text will highlight a few selected examples of socially inappropriate, violent or even grave criminal behaviour in people who have suffered brain damage, most commonly in the frontal lobe of the brain. First, the best-known textbook case regards the American railway constructor Phineas Gage, whose brain was injured in 1848 and his resulting behavioural and personality change has been described by his physician and subsequently quoted in many, if not most, textbooks of neurology, psychology and neuroscience. Another, more recent, line of consideration regards the findings that law offenders who are sentenced to death in the USA often suffer from unrecognized severe psychiatric, neurological, and cognitive disorders relevant to considerations of mitigation. Finally, aggressiveness reaching even homicidal levels may be a (rare) trait of infectious diseases of the nervous system, such as the Lyme disease. These examples raise many questions – to which degree may moral (or immoral) behaviour be influenced by brain dysfunction and ultimately whether the documented presence of brain damage and dysfunction may mitigate or excuse criminal conduct.

Keywords (Deutsch)
Hirnverletzung; Verhalten; erworbene Soziopathie; Frontallappen; Gewalt;

Keywords (English)
brain injury; behaviour; acquired sociopathy; frontal lobe; violence;

1. Introduction

Neurological and psychiatric literature since the early 19th century has presented case reports of patients with brain damage who manifested insensi-

tive, inappropriate, even aggressive and violent behaviour, typically when the lesion included the frontal lobe of the brain (Blumer & Benson 1982). The development of sociopathic behaviour after focal brain damage has been aptly termed "acquired sociopathy" (Tranel 1994). In the following text, we will select a few lines of evidence relevant for the topic.

First, the famous 19th century case of Phineas Gage will be described and discussed. The second perspective will take an opposite approach: Start from serious criminal behaviour and discuss the possible causal effect of brain damage on the frequently found psychiatric, neurological, and cognitive disorders relevant to considerations of mitigation. Finally, the occurrence of aggressive and violent behaviour within the clinical presentation of infectious diseases of the nervous system will be discussed.

2. The Textbook Brain Injury Case of Phineas Gage

One prominent 19th century case of head and brain injury, that of Phineas Gage, has become a textbook example of brain damage causing behavioural and personality changes, deserving a place in article titles like "Phineas Gage and the beginnings of neuropsychology" (Larner & Leach 2002).

Since the readers coming from different fields may not be familiar with the case, let me recount the basic facts about the man, his injury, and its consequences.

Phineas Gage was 25 years old at the time of the injury, literate, healthy, and strong, judged by his employers to be very efficient and capable, smart, energetic, and persistent in fulfilling his assigned job duties. On September 13, 1848, Gage was working as a railroad foreman, excavating rock with blasting powder for railroad construction in Vermont, USA. An accidental explosion thrust an iron bar called tamping iron (about 1 m long and 3 cm in diameter) up through Gage's head, penetrating his face below the cheekbone and exiting from the top of his skull. The iron landed some 25 meters away. Gage was thrown on his back and manifested some brief convulsions of all four extremities but got up and spoke within a few minutes, walked with little assistance and sat upright in an oxcart during a 1.2 km ride to his lodgings in the nearby village. Despite becoming blind

in the left eye and suffering from subsequent serious wound infection, Gage survived. Later research and recent three-dimensional reconstruction of his preserved skull showed that the tamping iron likely destroyed most of his brain's left frontal lobe.

After several months, Gage achieved, in some respects, full recovery. His physician since before the accident, John Martyn Harlow, not only saved his life but also wrote two papers about Gage, reporting both his physical condition and his behaviour before and after the injury. In the early 1848 paper (Harlow 1848), Gage regained his physical strength, manifested no deficit in movement or speech and was able to learn new information. His memory or intelligence seemed unaffected. On the other hand, Gage's once even-tempered personality changed dramatically. Harlow wrote "He is fitful, irreverent, indulging at times in the grossest profanity, which was not previously his custom." Gage's friends found him "no longer Gage". Apparently, he no longer showed the former balance of his "intellectual faculties and animal propensities". He would not keep plans and agreements, rather would change them at any moment, and for his fellows he had "little deference". In other words, he showed little of the former respect for social conventions and completely lost his sense of responsibility. Because of this deficit, the railroad-construction company that has employed him and had considered him a model foreman, refused to accept him back to work. Although Gage eventually secured employment, he never reached his former position and eventually joined relatives in San Francisco, where he died in May 1860, at age 36, after a series of seizures, likely due to post-traumatic epilepsy.

Eight years after Gage's death and twenty years after the accident, John Harlow wrote the second paper on Gage (Harlow 1868). There, Harlow expertly correlated Gage's cognitive and behavioural changes with presumed partial damage of the frontal lobe of the brain (Harlow 1868). At the time of publication of this second paper, other cases have been presented to illustrate specific behavioural deficits after focal brain lesions, namely those of language, motor control, and perception (Broca 1865; Wernicke 1874). Localization of language functions in the brain, even though not accepted universally, became a topic of active research and discussion. The speculation of Harlow, however, that a circumscribed brain area might be

responsible for planning and execution of socially suitable and acceptable behaviour, as well as reasoning, was even more surprising. The medical environment of the time was apparently not ready for such far-reaching consideration. Reasoning and social behaviour were strongly linked to ethics and religion and so could not be the subject of medical inquiry and explanation.

There was, in fact, another reason for dismissing Harlow's argument: He only learned of Gage's death 5 years after it had occurred and there had been no autopsy. Harlow himself actually accomplished the best he could have at the time: He gained permission from Gage's family to exhume the body and recover the skull. The skull, as well as the personal tamping iron of Gage, have since been kept by the Warren Anatomical Medical Museum at Harvard University. Nevertheless, exact description of the brain lesion, as performed *post mortem* for the patients of Broca and Wernicke, was not possible at the time. Some critics would thus claim that the lesion actually affected Gage's language area or even the nearby "motor centres" or persisted in their opinion that there is no localization of function in the human brain.

One of the few, who came to Harlow's defence was David Ferrier, a British physiologist in a lecture delivered at the Royal College of Physicians of London (Ferrier 1878). Ferrier reviewed the previous anatomical reconstructions of the probable trajectory based on the recovered skull and reasoned, contrary to the above critics, that the lesion spared both motor and language centres and damaged the left frontal cortex and that such damage would explain Gage's behavioural disturbances, which he aptly called a "mental degradation" (Ferrier 1878).

The challenge of Gage's brain lesion localization was taken up in late 20th century by Damasio *et al.* (1994) with the help of CT scan, 3D reconstruction and fitting a normal brain to Gage's reconstructed skull. The authors carefully considered possible trajectories of the iron rod and concluded that the lesion involved both left and right prefrontal cortices in a pattern that, based on more recent well documented lesions, causes a deficit of rational decision making and the processing of emotions. The damaged areas included limited parts of both the left and right frontal lobe, namely, the anterior frontal pole (tip) as well as the grey matter of the

lower inner cortical surfaces overlying the orbits of the eyes, so-called orbitofrontal cortex. Besides the grey matter (cortex), the lesions included small parts of the white matter core of the frontal lobes, more extensively in the left hemisphere than the right. The language and motor centres on the outer (lateral) surface of the frontal lobe of the brain were not affected. Thus, the more than 100-year old conclusion of Ferrier (1878) was correct. Subsequent analyses of multiple patients with frontal lobe damage, not only due to traumatic injury but also stroke and tumours, have supported the hypothesis that the ventromedial frontal region is important for emotion-based decision-making in the social context and explicit aggression (Pirau & Lui 2021; Blumer & Benson 1982; Grafman *et al.* 1996). In contrast, dorsolateral (pre)frontal cortex (dlPFC) is implicated in other aspects of cognition, including handling extra-personal space, objects, calculation, and language (Fuster 2015). This part of the frontal lobe remained intact in Gage, explaining his preserved functioning in many "rational" behavioural domains; in modern Gage-like patients, traditional neuropsychological testing of these cognitive domains may likewise show normal functioning. Interestingly, the dorsolateral prefrontal cortex, which is key for response inhibition, may also modulate aggression and violence but differently; lesions of the dlPFC in Vietnam War veterans were associated with more positive implicit attitude toward aggression and violence that would under normal conditions be considered inappropriate (Cristofori *et al.* 2016).

Back to Gage and his post-injury personality and behaviour. It is fair to note that Harlow's highly informative but still relatively concise statements on this matter have since been often generalized and exaggerated in some of the modern books, whose authors have apparently added their own impressions and filled in information not present in the original sources. Malcolm Macmillan from the University of Melbourne describes the later view of Gage in these words: "Phineas as an unstable, impatient, foulmouthed, work-shy drunken wastrel, who drifted around circuses and fairgrounds, unable to look after himself, and dying penniless in an institution" (Macmillan & Lena 2010, 643). Macmillan and Lena have evaluated all available information on Gage (Macmillan 2002) and called their subsequent paper "Rehabilitating Phineas Gage". Their summary emphasizes

Gage's significant psycho-social recovery; contrary to the above-mentioned distorted opinion, Gage

> worked and supported himself throughout his post-accident life; his work as a stage-coach driver was in a highly structured environment in which clear sequences of tasks were required of him; within that environment contingencies requiring foresight and planning arose daily; and medical evidence points to his being mentally unimpaired not later than the last years of his life. Although that Phineas may not have been the Gage he once had been, he seems to have come much closer to being so than is commonly believed. (Macmillan & Lena 2010, 655)

Thus, in addition to providing a textbook example of the relationship of brain damage, personality, and behavioural change, the case of Phineas Gage also puts forward clear evidence of functional recovery of brain functions after injury.

3. Serious Crime and Brain Damage

Since the 1980s, several reports have appeared describing high rates of psychiatric, neurological, and cognitive disorders among people convicted of serious crimes and sentenced to death (in the USA) (Lewis *et al.* 1986; Lewis *et al.* 1988; Martell 1992; Witzel *et al.* 2016). The obvious moral as well as legal question followed: Do these findings mitigate or excuse criminal conduct?

In 1986, Lewis and Pincus published a study of 15 death row inmates that found that all these inmates had suffered severe head injuries in childhood and about half had been injured by assaults. A comprehensive and uniform examination protocol was, unfortunately, not performed across all subjects, because of administrative obstacles and time constraints. Besides incomplete clinical examinations, only 5 had EEGs and 3 had head CT scans. Still, the evidence of head injuries was clear in all of them and 6 subjects were chronically psychotic. It is also necessary to state that the subjects were selected because of the imminence of their executions and

not because of obvious psychopathology (as the courts found no grounds for the insanity defence); thus, they may be considered representative of the criminals awaiting execution in the USA.

Furthermore, not only that the subjects themselves did not invoke any "abuse excuse" during the criminal process, all but one have actually minimized or denied their psychiatric disorders, apparently from the position that it was better to be "bad" than "crazy". One of the apparent results of their trauma, in fact, was limited recall of the details of their abuse. The clinical history has mostly been collected from childhood medical records and interviews with family members.

In the 1988 study by the same authors of 14 juveniles sentenced to death, all had suffered head trauma, most in car accidents but many by beatings as well. Twelve had suffered brutal physical abuse, 5 of those sodomized by relatives. In this age group, this would have a reasonable chance to serve for purposes of mitigation, however, for a variety of reasons, the subjects' injuries were not recognized at the time of trial and sentencing. In this study group, the examination protocol was more successfully realized, e.g., standard neurological examination was possible in 12 out of the 14 subjects, neuropsychological test battery was obtained in all and a neurometric quantitative EEG was performed in all subjects, although no CT scans were reported.

As mentioned above, the frontal lobe of the brain is the most commonly considered structure, its putative dysfunction has been cited to explain the acts of at least some people engaged in violent criminal behaviour. These apparently fail to inhibit impulsive, inappropriate, or habitual aggression (Brower & Price 2001). The clinical (neurological, psychiatric, psychological) deficits in law offenders were also accompanied and supported by imaging (CT, MRI, PET) findings documenting structural and functional changes of the brain. Frontal and superior parietal lobe cortical as well as multifocal subcortical (amygdala, thalamus) metabolic abnormalities (lowered glucose metabolism) were documented by positron emission tomography (PET) (Raine *et al.* 1997). The significance of these findings has been well-summarised by an article in the New York Times (Mansnerus 2001). Here, the journalist shares Dr. Lewis's broader perspective expressed in her book *Guilty by Reason of Insanity* (1998) that while no

revolution is at hand in the criminal justice system, legal scholars say new findings on brain dysfunction are finally gaining attention, at least where they matter most: in death penalty cases. Just this year, 4 states banned executions of the mentally retarded, bringing to 17 the number of the 38 death-penalty states that have made that exception, and the Supreme Court will hear arguments in one such case this fall. While no one would suggest that abuse or brain damage makes a murderer, Lewis argues in the aforementioned book that while most damaged people do not turn into killers, almost every killer is a damaged person. She concludes that most murderers are shaped by the combination of damage to the brain, particularly to the frontal lobes, which control aggression and impulsiveness, and the even more complex damage visited by repeated, violent child abuse. These findings, Lewis continues, cast doubt on legal definitions of insanity (Lewis 1998).

Many legal experts agree with Lewis, while others argue that the law should be in no hurry to apply new theories in the debate that is older than Western thought itself; namely, the one between free will and determinism. Many psychiatrists and psychologists, too, see evil and con artistry where researchers like Lewis see disease.

Barbara R. Kirwin, a forensic psychologist who recounted her examinations of violent murderers in her book, *The Mad, the Bad and the Innocent*, questions Lewis's studies because, like many medical studies with small samples, they are not controlled. And if unusual brain activity can be interpreted, Kirwin stated, "I want to find out what subcortical firing Mother Theresa has." (Kirwin 1997)

Kirwin's findings on the incidence of child abuse among homicide defendants differ wildly from Lewis's. Kirwin estimates that of the 300 or so defendants she has studied, 10 % have been abused, or "about what you'd find in the general population" (Kirwin 1997). One way of stating their differences is that Lewis claims she has never seen a "mere sociopath" – that is, someone with a normally competent brain who simply has a gross lack of empathy – while Kirwin concludes that she has seen plenty.

While the quoted side-by-side presentation by the New York Times writer of the contrasting views of Dr. Lewis and Ms. Kirwin may leave the reader undecided, a scientific review of Mr. Kirwin's book (O'Regan 1999)

reports that the book lacks a scientific rigor, relying almost exclusively on the author's personal (albeit professional) experience to support her arguments. Even when she inserts other, generally accepted, facts, no clues are provided to the reader to distinguish supported facts and personal opinions. The scientific publication record of Ms. Kirwin is minimal. All this suggests that the opposition to Dr. Lewis's perspective lacks solid ground and Dr. Lewis's concerns still deserve serious consideration.

3.1 Crime and Brain: A Case Study

Greely describes a highly relevant example for the relationship of altered brain and inclination to crime, as reported in the *Archives of Neurology* (Burns & Swerdlow 2003). A middle-aged man has suddenly developed an interest in child pornography. Shortly thereafter, he molested his 12-year-old stepdaughter, for which he was arrested and convicted. As a first-time offender, he was sent to a diversion program, but he failed the diversion program because he propositioned everyone he saw. Thus, he was scheduled to appear in court to be sentenced to prison. The day before the scheduled court appearance, a severe headache prompted him to visit an emergency room. He was admitted by the psychiatry service because an organic cause of his headache was not suspected at the time. An MRI scan performed later revealed a benign tumour in his frontal lobe, several centimetres in size. After surgical removal of the tumour, the man reportedly lost all interest in pornography. He took the diversion program again and this time passed easily. He was therefore not sent to prison, but attempted to rebuild his life. About a year later, the headache returned and the man again began secretly gathering pornography. A new brain CT scan showed recurrence of the tumour. It was once again surgically removed and, again, the disturbing sexual impulses disappeared.

In his paper, professor Greely presents several disconcerting questions to the reader:

> If that's the defendant in front of you – as a prosecutor, a judge, or a parole board–what do you do with him? And why? This was not a neuroscience case; it was a case where the "external cause", if a tumour inside one's own

> skull can be called "external", is extraordinarily, though still not perfectly, clear. But I suspect neuroscience will give us more such cases, either in rare individuals or in unusual classes of people. If so, the law will have to decide how to handle such offenders. (Greely 2015, 700–701)

These questions may be easily rephrased in moral, not legal, terms. And finding the right moral answers may be as challenging as the legal ones.

4. Nontraumatic Brain Damage and Behaviour: The Case of Lyme Disease (Neuroborreliosis)

Nontraumatic brain damage may also manifest with behavioural changes. Among infectious diseases, one not so commonly considered cause is Lyme disease (LD), a multisystem infection which initially affects the skin but may spread to joints, the heart, and the nervous system (about 10–15 %). Within the nervous system, both the peripheral (peripheral nerves and ganglia) and the central (brain and spinal cord) divisions may be affected. Recently, a psychiatrist reported first increased suicidality (Bransfield 2017) and later also homicidality (Bransfield 2018) in patients suffering from Lyme disease.

In the 2018 paper, Bransfield reports that retrospective analysis of 1000 LD patient charts in psychiatric care found that about 10 % of these patients were homicidal, with the average diagnosis delay of 9 years (Bransfield 2018, 693). Aggression in psychiatric LD was impulsive, sometimes provoked by intrusive symptoms, sensory stimulation, or frustration and invariably bizarre and senseless. (Bransfield 2018, 693) Comparing the psychiatric profiles of a homicidal LD subgroup with a matched non-homicidal LD subgroup yielded association of other psychiatric disturbances in the homicidal group, including suicidality, abrupt mood swings, explosive anger, paranoia, anhedonia, hypervigilance, exaggerated startle, disinhibition, nightmares, depersonalization, intrusive aggressive images, dissociative episodes, derealization, intrusive sexual images, substance abuse, depression, panic disorder, memory impairments, and decreased libido as well as neurological disturbances – neuropathy and cranial nerve symptoms" (Bransfield 2018, 693).

The author of the aforementioned study concludes that while many LD patients have no aggressive tendencies or only mild degrees of low frustration tolerance and irritability and pose no danger, a lesser number experience explosive anger, a yet lesser number experience homicidal thoughts and impulses, and much lesser number commit homicides. (Bransfield 2018, 693)

While LD affects less than 1⁄1000 people a year, less than 10 % of them develop late-stage nervous system affection and out of those, again a minority seeks psychiatric care. Nevertheless, the observations of Bransfield (2018) serve as a reminder that in a very small minority of patients, the behavioural disorder can be highly significant, threatening the life of the patient himself or the people around him.

5. Conclusion

It has long been known that brain injuries and lesions may influence human behaviour. We have reviewed a few selected examples of socially inappropriate, violent or even grave criminal behaviour in people who have suffered brain damage, most commonly in the frontal lobe of the brain. These examples raise many questions – to which degree may moral (or immoral) behaviour be influenced by brain dysfunction and ultimately whether the documented presence of brain damage and dysfunction may mitigate or excuse criminal or immoral conduct.

Bibliography

Blumer, D., Benson, D. F., Personality Changes with Frontal and Temporal Lobe Lesions, in: Blumer, D., Benson, D. F. (eds), Psychiatric Aspects of Neurological Disease, New York, NY 1982, 151–170.

Bransfield, R. C., Suicide and Lyme and Associated Diseases, "Neuropsychiatric Disease and Treatment", 2017, 13, 1575–1587, https://doi.org/10.2147/NDT.S136137.

Bransfield, R. C., Aggressiveness, Violence, Homicidality, Homicide, and Lyme Disease, "Neuropsychiatric Disease and Treatment", 2018, 14, 693–713, https://doi.org/10.2147/NDT.S155143.

Broca, P., Sur le siège de la faculté du langage articulé, "Bulletins et Mémoires de la Société d'Anthropologie de Paris", 1865, 6, 5, 377–393.

Brower, M. C., Price, B. H., Neuropsychiatry of Frontal Lobe Dysfunction in Violent and Criminal Behaviour. A Critical Review, "Journal of Neurology, Neurosurgery, and Psychiatry", 2001, 71, 6, 720–726, https://doi.org/10.1136/jnnp.71.6.720.

Burns, J. M., Swerdlow, R. H., Right Orbitofrontal Tumor with Pedophilia Symptom and Constructional Apraxia Sign, "Archives of Neurology", 2003, 60, 3, 437–40, https://doi.org/10.1001/archneur.60.3.437.

Cristofori, I. *et al.*, Brain Regions Influencing Implicit Violent Attitudes. A Lesion-Mapping Study, "The Journal of Neuroscience: The Official Journal of the Society for Neuroscience", 2016, 36, 9, 2757–2568, https://doi.org/10.1523/JNEUROSCI.2975-15.2016.

Damasio, H. *et al.*, The Return of Phineas Gage. Clues about the Brain from the Skull of a Famous Patient, "Science", 1994, 264, 5162, 1102–1105, https://doi.org/10.1126/science.8178168.

Ferrier, D., The Goulstonian Lectures on the Localisation of Cerebral Disease I. Conclusion, "British Medical Journal", 1878, 1, 900, 443–447, https://doi.org/10.1136/bmj.1.900.443.

Fuster, J., The Prefrontal Cortex, Amsterdam ^{5}2015, https://www.elsevier.com/books/the-prefrontal-cortex/fuster/978-0-12-407815-4.

Grafman, J. *et al.*, Frontal Lobe Injuries, Violence, and Aggression. A Report of the Vietnam Head Injury Study, "Neurology", 1996, 46, 5, 1231–1238, https://doi.org/10.1212/wnl.46.5.1231.

Greely, H., Law and the Revolution in Neuroscience. An Early Look at the Field, "Akron Law Review", 2015, 42, 3, 687–715, https://ideaexchange.uakron.edu/akronlawreview/vol42/iss3/2.

Harlow, J. M., Passage of an Iron Rod through the Head, "The Boston Medical and Surgical Journal", 1848, 39, 389–393, https://doi.org/ 10.1056/NEJM184812130392001.

Harlow, J. M., Recovery from the Passage of an Iron Bar through the Head, "Publications Mass Med Soc", 1868, 2, 327–347.

Kirwin, B. R., The Mad, the Bad, and the Innocent. The Criminal Mind on Trial, Boston, MA 1997.

Larner, A., Leach, J. P., Phineas Gage and the Beginnings of Neuropsychology, "Advances in Clinical Neuroscience and Rehabilitation", 2002, 2, 3, 26.

Lewis, D. O. *et al.*, Neuropsychiatric, Psychoeducational, and Family Characteristics of 14 Juveniles Condemned to Death in the United States, "The American Journal of Psychiatry", 1988, 145, 5, 584–89, https://doi.org/10.1176/ajp.145.5.584.

Lewis, D. O. *et al.*, Psychiatric, Neurological, and Psychoeducational Characteristics of 15 Death Row Inmates in the United States, "The American Journal of Psychiatry", 1986, 143, 7, 838–845, https://doi.org/10.1176/ajp.143.7.838.

Lewis, D. O., Guilty by Reason of Insanity. A Psychiatrist Probes the Mind of Killers, New York, NY 1998.

Macmillan, M., An Odd Kind of Fame. Stories of Phineas Gage, Cambridge, MA 2022, https://books.google.cz/books?id=Qx4fMsTqGFYC.

Macmillan, M., Lena, M. L., Rehabilitating Phineas Gage, "Neuropsychological Rehabilitation", 2010, 20, 5, 641–58, https://doi.org/10.1080/09602011003760527.

Mansnerus, L., Damaged Brains and the Death Penalty, The New York Times, 21 July 2001, section Arts, https://www.nytimes.com/2001/07/21/arts/damaged-brains-and-the-death-penalty.html.

Martell, D. A., Estimating the Prevalence of Organic Brain Dysfunction in Maximum-Security Forensic Psychiatric Patients, "Journal of Forensic Sciences", 1992, 37, 3, 878–93.

Matuszewski, A., The Mistaken Emphasis on Organic Brain Damage in Capital Habeas Litigation, "Alabama Law Review", 2015, 67, 1217.

O'Regan, M., The Mad, the Bad, and the Innocent. The Criminal Mind on Trial, "Psychiatric Services", 1999, 50, 1, 125–126, https://doi.org/10.1176/ps.50.1.125.

Pirau, L., Lui, F., Frontal Lobe Syndrome, in: Abai, B. *et al.* (eds), StatPearls, Treasure Island, FL 2021, https://www.ncbi.nlm.nih.gov/books/NBK532981/.

Raine, A. *et al.*, Brain Abnormalities in Murderers Indicated by Positron Emission Tomography, "Biological Psychiatry", 1997, 42, 6, 495–508, https://doi.org/10.1016/S0006-3223(96)00362-9.

Tranel, D., Acquired Sociopathy. The Development of Sociopathic Behavior Following Focal Brain Damage, "Progress in Experimental Personality & Psychopathology Research", 1994, 17, 285–311.

Wernicke, C., Der aphasische Symptomencomplex. Eine psychologische Studie auf anatomischer Basis, Breslau 1874.

Witzel, J. G. *et al.*, Increased Frequency of Brain Pathology in Inmates of a High-Security Forensic Institution. A Qualitative CT and MRI Scan Study, "European Archives of Psychiatry and Clinical Neuroscience", 2016, 266, 6, 533–541, https://doi.org/10.1007/s00406-015-0620-2.

Field Reports

Erfahrungsberichte

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 113– 119
DOI: 10.24989/BCE.neuro.7

Psychotic Disorders and Personal Freedom

Psychotische Störungen und personale Freiheit

Borut Škodlar, Ljubljana

Abstract (Deutsch)
In diesem Bericht geht der Autor der Frage nach, ob persönliche Freiheit für Menschen, die mit psychotischen Störungen zu kämpfen haben, überhaupt als Realität denkbar ist. Er untersucht diese Frage anhand einer Fallstudie über Schizophrenie. Im ersten Teil beschreibt er, welche Herausforderungen die Schizophrenie für Patient*innen darstellt, die mit ihr zu kämpfen haben, und wie sie die Entscheidungsprozesse der Patient*innen beeinflusst. Diese Beschreibung untermauert er dann im zweiten Teil mit zwei Berichten seiner eigenen Patient*innen, die sich zum Zeitpunkt der Behandlung in unterschiedlichen Stadien der Entwicklung der Störung befanden. Im letzten Teil des Berichts stellt der Autor eine kurze Reflexion über den Fall von John T. Perceval an, um zu zeigen, dass das Leben mit einer psychotischen Störung nicht zwangsläufig mit einem vollständigen Verlust der persönlichen Freiheit einhergeht.

Abstract (English)
In this report, the author tackles the issue of whether personal freedom is at all thinkable as a reality for those people who are struggling with psychotic disorders. He investigates the issue by using a case study of schizophrenia. In the first part, he describes the kind of challenges schizophrenia poses to a patient struggling with it and how it affects the patient's decision-making processes. He then substantiates this description in the second part with two accounts of his own patients who were at the time of treatment at different stages of the development of the disorder. In the last part of the

report, the author offers a short reflection on the case of John T. Perceval in order to argue that living with a psychotic disorder does not necessarily imply a complete loss of personal freedom.

Keywords (Deutsch)
Psychose; Schizophrenie; Psychotherapie; Freiheit; Resilienz;

Keywords (English)
psychosis; schizophrenia; psychotherapy; freedom; resilience;

1. Introduction: Schizophrenia and the Limits of Freedom

I will shortly describe my view on the subject through the lens of psychotic disorder, more specifically schizophrenia. Both very traumatised people – be it victims of individual or collective trauma – as well as individuals within schizophrenia spectrum from prodromal stages and schizotypy to the full-blown schizophrenia have difficulties with building and constructing their autobiographies, their own stories of personal lives due to their specific problems.

Schizophrenia is one of the most severe mental disorders and causes large health expenditure. Being one of the paradigmatic psychotic disorders, schizophrenia is also relatively well researched. It is a chronic mental disorder or brain disorder, as it could be also conceived. It is also a uniquely human illness. What I would pinpoint from neurobiological research is one aspect; namely, that elevated dopamine in some parts of the brain is causally related, as it seems and as the research shows, to what is called aberrant salience as one of the researchers put it (Kapur 2003). Aberrant salience means that people experience outer stimuli, outer sensory data, including intersubjective, interpersonal data, too intensively. People with schizophrenia perceive the outer events as well as inner thoughts and experiences as very important, relevant for them or salient. The threshold for salience is in people vulnerable to schizophrenia very low, and they thus experience as relevant too many impulses and impressions from the outer world. In other words, in the normal course of a day, we filter out a lot

of data, impressions, and sensations as not relevant and important. We simply store them into the tacit, background dimension of our experience, and we can focus on, for example, having a conversation or on any other thing that we choose to focus on. Meanwhile, patients with schizophrenia are unable to filter out, to set aside sensory data, which leads to them being overwhelmed constantly by such sensory data. That is why they cannot cope with that in intersubjective and other situations; they simply freeze, or they feel like they are paralysed and totally inhibited and that is also the reason why they often withdraw from social interactions.

The main phenomenological research on schizophrenia shows exactly the same; the patients tell us that they cannot ward off or shelter themselves from the outside world, especially from other people, and they are constantly overwhelmed by the presence of them. They feel overwhelmed by the constant threat of the gaze of other people, frequently frozen in social situations and they can thus not be with others or focus on the interpersonal relationships. At the same time, they cannot focus on their activities and pursue their life goals. So, they constantly struggle in their everyday life and feel defeated by these struggles (Sass 2008).

2. Two Clinical Vignettes

In the continuation, I will narrate two short vignettes that illustrate what I was now describing. Then I will come to the point of what can people vulnerable to schizophrenia do about it, how free they are to choose, construct, and lead their lives. I will proceed in this via an example of an important historical figure who showed through his life the struggle of how to cope with schizophrenia.

So, first to two short vignettes. They refer to two patients who were treated in our Unit for Psychotherapy at the University Psychiatric Clinic Ljubljana (Slovenia). One is a 19-year-old patient, who came to our unit in the so-called prodromal phase of psychotic disorder, which means he has never experienced a full-blown psychosis. He was not yet diagnosed with schizophrenia, but the clinical picture of his prodromal experiences was indicating the schizophrenia spectrum. What was he experiencing,

and why did we think that he was most probably suffering from emerging schizophrenia? The main reason was his feeling of being constantly overwhelmed by any interaction with another human being. He felt overwhelmed by those interactions so much, that he felt others were using him, manipulating him, and his threshold for feeling that way was very, very low. Of course, we can all think sometimes that somebody is perhaps manipulating us, but he felt that way with almost everybody, including his parents, his sister, his close friends. He felt this world is a cold, manipulative place not worth living and a nihilistic worldview was a natural outcome of such experiences and thought processes. He was admitted to our unit, but soon, i.e., in a week, "escaped" from it, because he could not cope with being with others in the unit, feeling that they were manipulative and abusive. So, he went home and not long afterwards he attempted suicide. Luckily, he was rescued and brought to the hospital again, and now he is seeing me as an outpatient and is more open than before to share his experiences and also to seek help; even though he is not very optimistic and not spontaneously pervaded by hope that he can be helped he nonetheless keeps coming and our therapy process progresses slowly. This was a patient in an initial stage of schizophrenia.

Another patient has experienced four psychotic episodes already and is currently 37 years old. He has undergone, as I said, several severe psychotic episodes and is at the moment in a relatively good remission. However, he also has great difficulties relating to others. He lives with his parents, has almost no social interactions; just one friend whom he sees once a week or once every second week, and a few, sporadic phone calls with some colleagues, but nothing more than that. Hence, he lives in such an isolated environment and he is constantly fearful that some evil forces will take him, that he will lose control over his mental capacities, his identity, and awareness of who he is. These are actually fears from his psychotic episodes during which he felt that he was overtaken by the evil forces and they continue to influence his everyday life. At this stage, he thus feels that his inner life and inner situation are somehow unprotected from evil forces that are often represented by other people. Both patients, therefore, live in the same constant fear or uneasiness about being with other people. The sense of defencelessness – so called *schiz-*

ophrene Wehrlosigkeit in German psychopathology (Burkhart 1962) – is their constant companion.

3. Free choice and resilience in schizophrenia: the case of John T. Perceval

This raises the question: What is and to what extent do these patients have freedom and free choice to lead a life worth living, as it were. Viktor Frankl, the founder of logotherapy, is known for his emphasis on the importance of meaning in one's life, claimed that there is a degree of freedom in every human situation; even in such constricted situations that our patients go through. But, what kind of freedom is possible in that state and in those situations? As a way of answering that question, I will present a very interesting historical case, discovered by a famous anthropologist, Gregory Bateson, in the 1960s (Bateson 1961). The case concerns John Thomas Perceval, a British army officer and a member of the British nobility who lived between 1803 and 1876 and whose autobiographical notes were discovered and subsequently published by Bateson. During the course of his life, Perceval was confined to two lunatic asylums of those times. He was from a well-off family; his father Spencer Perceval was a prime minister, but shot to death when John Perceval was nine years old. After growing up, he had a brilliant career as an army officer, but, at some point, felt that the army was not his place where he would feel at home. Consequently, he looked for something else, first enrolling into university studies and, later on, embarking on a spiritual quest during which he came across and joined a radical evangelical sect. The sect was engaged in fringe religious practices such as speaking in tongues and performing miracles, while claiming for itself the guidance of the Holy Spirit. Perceval's spiritual quest, however, somehow took a psychotic turn. He entered a full-blown psychosis and was taken by his family to a lunatic asylum, where he felt completely alone and abandoned. These old asylums, one has to keep in mind, engaged in quite severe and brutal therapeutic procedures, such as administering cold baths and performing various surgeries on patients – and Perceval was not spared any of these "therapies" on account of his family background. This first episode happened in 1830, when

Perceval was 27 years old. According to his testimony, in that asylum Perceval realised that he needed to distance himself from delusions and in order to be able to do that, he would need to impose on himself a very strict regime of discipline. Apart from that, he constantly strived to recollect himself, to be fully present so as to be able to get back to his everyday life. He felt that he needed to synchronise body and mind in order to avoid psychotic episodes, and he identified strengthening daily routines and a very disciplined way of life as means to that end. This allowed him not to be distracted and to be able to focus, which presents, as I mentioned earlier, a tremendous challenge for such patients. Through such efforts that indicated his amazing resilience and mental struggle, *Trotzmacht des Geistes* as Viktor Frankl would call it (Frankl 1997), he found a way how to cope with his illness. Subsequently, Perceval convinced the psychiatrists at the asylum that he was doing well. He convinced his family to get him out of the asylum, and he devoted his life after that to taking care of inmates of asylums. He also founded an organisation called *Alleged Lunatics' Friend Society*, which was the first organisation in the world to fight for the rights of the inmates, of psychotic, schizophrenic, and other patients brought to asylums or elsewhere.

4. Conclusion: Therapeutic Freedom and Optimism

With an example like Perceval, which is well documented through his notes, we can also help our patients as they struggle with their illnesses, so as to find ways how to cope with them.

Perceval's notes show that he found ways to use his freedom to cope with his illness:

> Whenever my thoughts and hands were most occupied, I became, I suppose, nearest to a sound state of mind, and consequently more aware of my situation, that all or that many of the faculties of mind and body should be called into play at one time, and above all things that the body should be occupied,

and he concluded by stating: "My soul survived that ruin." (Perceval 1840, 285).

Similarly, as Perceval's personal experience show us, we can also encourage our patients, that there are at least some aspects of freedom in every possible life situation and that people can always activate it even if they suffer from a very severe and long-lasting mental disorder, such as schizophrenia. The axiom of freedom and psychiatric *credo* must survive through all the struggles of doubt and therapeutic pessimism be it in patients or in us therapists as Viktor Frankl always stressed, and we are called to bring this vision to our suffering fellow human beings (Frankl 1997).

Bibliography

Bateson, G. (ed.), Perceval's Narrative. A Patient's Account of His Psychosis 1830–1832, Stanford 1961.

Burkhardt, H., Die schizophrene Wehrlosigkeit, "Nervenarzt", 1962, 33, 306–312.

Frankl, V. E., Ärztliche Seelsorge. Grundlagen der Logotherapie und Existenzanalyse, Berlin 1997.

Kapur, S., Psychosis as a State of Aberrant Salience. A Framework Linking Biology, Phenomenology, and Pharmacology in Schizophrenia, "American Journal of Psychiatry", 2003, 160, 13– 23.

Perceval, J., A Narrative of the Treatment Experienced by a Gentleman, During a State of Mental Derangement, London 1840.

Sass, L. A., Madness and Modernism. Insanity in the Light of Modern Art, Literature, and Thought, Oxford 2017.

Bioethik in der Diskussion 4 (2022)
The Moral Brain
S. 121–124
DOI: 10.24989/BCE.neuro.8

Authors

Prof. Dr. Elisabeth Hildt

is a Professor of Philosophy and the Director of the Center for the Study of Ethics in the Professions at the Illinois Institute of Technology in Chicago. Her research focuses on bioethics, ethics & technology, research ethics, and ethical aspects of information and communications technology and artificial intelligence.

Prof. Dr. Petr Hluštík, M.D.

is a professor in the Department of Neurology, Faculty of Medicine and Dentistry, Palacký University, Olomouc, Czech Republic. His work combines teaching neurology and neuroscience courses at both undergraduate and postgraduate levels at Palacký University, supervising Ph.D. students, research in functional mapping brain including sensorimotor system plasticity and cognitive studies of language, memory, emotion and music perception in health and disease and clinical duties at the Faculty Hospital Olomouc. Prof. Hluštík trained in Brno and Olomouc (Czech Republic) and in Pittsburgh, Baltimore and Chicago (USA).

Prof. Dr. Piotr Jan Morciniec

is professor of bioethics and moral theology. In 2000, he was awarded the Individual Prize of the Minister for Education for his habilitation thesis "Ethical Aspects of Transplantation Therapy for Neurodegenerative Diseases". From 2010–2019, he was head of the Institute for Family Studies at the University of Opole. Since 2019, he is head of the Chair of Moral Theology, Bioethics and Canon Law. He is a member of the European Society for Catholic Theology and the Association of Bioethics in Central Europe; a member of the Expert Team on Bioethics of the Polish Conference of Bishops. He is the chief editor of the international academic journal "Family Forum". One of his most important publications of the past years is the book *Personalistic Bioethics with Regard to Dealing with the Human Corpse* (Opole 2009).

Univ.-Prof. Dr. Sigrid Müller

is professor of moral theology and teaches ethics at the Catholic Theological Faculty of the University of Vienna. She is the co-founder of the Association of Bioethicists in Central Europe and the CEEPUS network "bioethics". She representative the Catholic Theological Faculty as deputy chair of the interdisciplinary Institute for Ethics and Law in Medicine, which is a joint venture of the University of Vienna and the Medical University of Vienna.

Univ.-Prof. Dr. Nenad Polgar

is a professor of moral theology at the Faculty of Catholic Theology of the Karl Franzens University of Graz. His research interests and publication foci are: theological anthropology, moral methodology, sexual ethics, and bioethics. He recently completed his habilitation thesis and is currently preparing two monographs for publication: *Understanding Homosexuality* and *The Meanings, Origins, and Relevance of the Concept of Intrinsic Evil*.

He is a member of the Young Curatorium of the European Society for Catholic Theology.

Prof. Dr. Borut Škodlar

is Head of the Unity for Psychotherapy at the University Psychiatric Hospital Ljubljana and Associate Professor of Psychiatry at the Medical Faculty, University of Ljubljana, Slovenia. Among his main interests are phenomenological research and psychotherapy of psychotic disorders. Within psychotherapy, his main focus is psychoanalytical and existential psychotherapy in relation to contemplative practices, such as meditation and prayer. Within phenomenological research, he focuses on intricacies between psychotic and mystical states, and on the value of mystical and religious experiences in the recovery process as well as on suicidality in severe mental disorders.

Univ.-Prof. i. R. DDr. Walter Schaupp

studied medicine and theology. Since 2003 he was professor of moral theology at the Karl Franzens University of Graz, Faculty of Catholic Theology, where he retired in 2019. His main areas of work are medical ethics and bioethics as well as ethics of the good life and ethics and spirituality.

Prof. Dr. Andrea Vicini, SJ

is Chair of the Theology Department, Michael P. Walsh Professor of Bioethics, and Professor of Theological Ethics at Boston College (USA). He is an alumnus of Boston College (S.T.L. and Ph.D.). He also holds an M.D. and a specialization in pediatrics from the University of Bologna and an S.T.D. from the Pontifical Faculty of Theology of Southern Italy in Naples (Italy). At Boston College, he was Gasson Professor (2009–2010) and taught at the School of Theology and Ministry (2011–2019).

He also taught in Italy, Albania, Mexico, Chad, and France. He is co-chair of the international network Catholic Theological Ethics in the World Church, as well as lecturer and member of associations of moral theologians and bioethicists in Italy, Europe, and the USA. His research interests and publications include theological bioethics, global public health, new biotechnologies, environmental issues, and fundamental theological ethics.

Univ.-Prof. Dr. Angelika Walser

lives and works in Vienna and Salzburg (Austria). She is married and has two daughters. Since 2015, she holds the Chair of Moral Theology at Salzburg University. Walser studied in Würzburg and Munich and finished her studies 1998 with a PhD in Moral Theology. From 1997–2005 she was teacher for religious education at various schools, in the adult education of the Archdiocese of Vienna and teaching assistant at different Austrian universities. 2007–2010 she received an APART-scholarship of the Austrian Academy of Sciences for the research project: *Questioning the Autonomy of Women in Areas of Bio-ethical Conflict at the Beginning of Life as an Investigation into Theological Ethics* (Habilitation 2013 at Vienna University). 2010–2013 she was platform ranger of the research platform "Religion and Transformation in Contemporary European Society" (RaT), 2013–2015 she was Professor at KPH Vienna-Krems (Teachers' education).